Corrado Mangione
Silvio Bozzi

STORIA DELLA LOGICA

Da Boole ai nostri giorni

Garzanti

CORRADO MANGIONE - SILVIO BOZZI

Storia della logica

Da Boole ai nostri giorni

GARZANTI

INDICE GENERALE

CAPITOLO TERZO

LA NASCITA DELLA LOGICA MATEMATICA

CAPITOLO QUARTO

LE ANTINOMIE E IL PROBLEMA DEI FONDAMENTI

CAPITOLO QUINTO

LA SVOLTA DEGLI ANNI TRENTA

Questo volume costituisce una versione riveduta e aggiornata di alcuni dei capitoli dedicati alla storia della logica comparsi a firma di Corrado Mangione nella *Storia del pensiero filosofico e scientifico* di Ludovico Geymonat, pubblicata fra il 1970 e il 1976 presso questo stesso editore. Sostanzialmente si tratta dei capitoli dedicati allo sviluppo della logica dall'Ottocento in poi ed è soprattutto nell'opera di aggiornamento e revisione che è intervenuto il secondo autore.

Dalla pubblicazione dell'ultimo volume (il VII) della prima versione (che era già un aggiornamento rispetto ai precedenti) sono passati più di quindici anni e se anche l'arco temporale coperto nella presente edizione non è essenzialmente diverso da quello allora trattato, la necessità di ampliamenti e modificazioni si giustifica sulla base del fatto che nel frattempo molte direttrici di ricerca si sono cristallizzate e sono emerse in modo più netto, sicché il tempo trascorso ha reso possibile e necessaria una visione più distaccata e ci sembra anche più articolata di tutta intera la storia. A questo si aggiunga la maggior disponibilità dei testi, non solo originali, ma anche di carattere espositivo, che ci sembrava rendesse necessaria una diversa lettura e presentazione di alcuni temi specifici.

Come ricordato, in questa nuova versione non sono stati sostanzialmente variati i limiti temporali e *grosso modo* la trattazione giunge fino ai primi anni settanta (anche se ragioni di completezza ci hanno a volte portato a superare questi limiti e ad accennare a sviluppi propri della ricerca degli ultimi anni). Non c'è una ragione «profonda» per questa limitazione temporale e sono state essenzialmente ragioni pratiche che ce l'hanno suggerita. È innegabile infatti che gli anni sessanta abbiano segnato una svolta nella ricerca logica: basta pensare alla creazione dell'Analisi non

standard, all'introduzione del *forcing*, alla soluzione del decimo problema di Hilbert, ecc. Al di là dei singoli risultati, è proprio in questi anni che si assiste ad un approfondimento dei legami fra ricerca logica e specifici settori dell'indagine matematica e filosofica, con un conseguente specializzarsi delle tecniche e delle stesse problematiche. Sarebbe stato quindi estremamente difficile per noi e per il lettore tentare una trattazione organica degli sviluppi dopo il 1970 senza correre il rischio di forzature, omissioni e fraintendimenti.

La ragione di fondo che ci ha spinto ad intraprendere questo lavoro è stata infatti solo quella di offrire al lettore – e a noi stessi – una visione complessiva delle problematiche che hanno portato alla ricerca di questi ultimi anni, analizzandone le origini e indagandone i legami e i presupposti. Se da una parte quindi – e questo riguarda soprattutto l'Ottocento – non abbiamo inteso compiere un'analisi puntuale nello spirito di una ricerca di carattere storiografico in senso stretto, dall'altra – e questo riguarda soprattutto gli ultimi due decenni – non abbiamo neanche voluto fornire una serie di aggiornamenti sullo stato della ricerca che non solo in molti casi non saremmo stati in grado di dare, ma che inoltre, a nostro avviso, non avrebbe neanche avuto senso tentare. Se la soluzione «di mezzo» che abbiamo adottato può essere di qualche utilità non siamo noi a doverlo dire, ma quello che ci sembra in ogni caso importante ribadire è l'urgenza – almeno sul piano didattico – di una visione panoramica di carattere storico che permetta a chi vuole interessarsi di logica, sia come profano, sia come futuro ricercatore, di metterne in prospettiva problematiche e metodi.

La ricerca logica – nel senso di ricerca logico-matematica – è relativamente recente e la perdita di una prospettiva storica ci sembrerebbe in questo caso particolarmente pericolosa in quanto favorirebbe, come già in parte ha fatto, una visione troppo riduttiva e troppo localizzata. Se ha un senso, almeno all'inizio della propria formazione di ricercatore, affrontare le questioni in un'ottica di *problem solving* all'interno di una disciplina in cui i nessi tra problemi e lo sfondo teorico sono comunque storicamente articolati, ci sembra che questo possa avere effetti negativi in un ambito di ricerca giovane, in cui il respiro corto, più che segno di slancio, può essere sintomo di scarsa vitalità: in altri termini, un conto è

occuparsi senza preoccupazione di nessi storici o filosofici di un problema di geometria, un altro invece è affrontare uno specialissimo problema logico che nella storia della disciplina ha ancora rare se non nulle diramazioni, senza preoccuparsi di conoscerne il contesto. Non ci pare un caso che in questi ultimi anni da più parti l'interesse per la storia della logica nel Novecento abbia avuto notevoli sviluppi e questo non solo da parte di storici, ma di logici impegnati nella ricerca. Questo sì a nostro avviso è segno di vitalità, in quanto testimonia la volontà di rivedere confini già stabiliti, presupposti dati per scontati, ecc.

Anche fissati i limiti dei primi anni settanta, ci sono però due esclusioni che dobbiamo esplicitamente giustificare e che solo a malincuore abbiamo operato. La prima riguarda l'approccio categoriale alla logica, più specificamente le ricerche iniziate da F. William Lawvere e dalla sua scuola, che appunto prendono più ampio respiro nei primi anni settanta. Nel settimo volume di aggiornamento dell'edizione originale, avevamo tentato una presentazione di questo indirizzo nella convinzione – che ancora condividiamo – che si tratti di uno dei passi più significativi della ricerca – non solo logica – di questi ultimi decenni. La decisione di sopprimere l'intero capitolo e di non sostituirlo con uno nuovo è dovuta essenzialmente al fatto che la ricerca nel frattempo ha avuto uno sviluppo estremamente rapido e che ci sarebbe stato impossibile darne un'idea corretta e adeguata entro limiti di spazio ragionevoli. D'altra parte, la trattazione che ne avevamo dato – seppure aveva fornito a non pochi studenti una prima indicazione di questo indirizzo di ricerca – non ci sembrava più adeguata oggi, dopo quindici anni. Quello che vogliamo ribadire è che si tratta sicuramente di indagini di estrema importanza, per una presentazione più adeguata delle quali siamo costretti a rimandare – oltre che ovviamente ai lavori originali – ai testi, ad esempio, di P.T. Johnstone, P. Scott e J. Lambek, R. Goldblatt, M. Barr e R. Wells citati nella bibliografia dell'ultimo capitolo.

Discorso diverso vale per l'informatica teorica o l'intelligenza artificiale, che è l'altro argomento oggi spesso abbinato alla logica e di cui noi non parliamo. In questo caso le ragioni temporali sono più giustificate nel senso che il grosso della ricerca ha trovato una sua organizzazione e diffusione verso la fine degli anni settanta ed è maturato pienamente solo negli anni ottanta. A questo proposito

ci siamo quindi limitati ad accenni ogniqualvolta i temi di cui stavamo parlando rendevano opportuno questo richiamo.

In queste decisioni hanno avuto un ruolo non trascurabile anche problemi riguardanti i prerequisiti richiesti a un eventuale lettore. In generale abbiamo presupposto da parte di chi legge, oltre ovviamente a un interesse non superficiale per la materia, un minimo di cultura filosofica e matematica ed in particolare, per quanto riguarda le conoscenze di logica, almeno una infarinatura corrispondente a un corso propedeutico di logica in facoltà di filosofia. Per questa ragione, mentre alcuni temi che supponiamo noti al lettore sono stati trattati senza discussioni dettagliate, siamo entrati più nei particolari nei casi in cui abbiamo ritenuto che il lettore non avesse conoscenze preliminari.

Due ultime osservazioni, di carattere più specifico. La prima riguarda la bibliografia. Esistono bibliografie generali di logica matematica, basti pensare a quella classica di Alonzo Church del 1936, successivamente aggiornata, e a quella più recente a cura di Gert H. Müller e collaboratori (1987, in sei volumi). Questo ci ha confermato nell'opportunità di non tentare nemmeno di dare una bibliografia con aspirazioni non diciamo alla completezza, ma anche solo effettivamente rappresentativa. Abbiamo quindi deciso di limitarci a brevi indicazioni capitolo per capitolo, privilegiando, ove possibile, raccolte, atti di convegni, ecc.; in generale, la maggior parte dei volumi citati contiene dettagliate bibliografie, sia per quanto riguarda i testi che gli studi: è per questo motivo che, salvo per casi molto particolari, abbiamo evitato di citare articoli la cui elencazione avrebbe immediatamente appesantito in modo sproporzionato la stessa informazione bibliografica. D'altra parte, molti lavori vengono citati direttamente nel testo; in questo caso, in particolare per quanto riguarda volumi (piuttosto che articoli) ci siamo limitati a riportare in bibliografia le eventuali traduzioni italiane degli stessi. In generale le indicazioni bibliografiche seguono l'ordine dei paragrafi del capitolo.

Nella stessa ottica ci siamo mossi per quanto riguarda un secondo problema, che diventa particolarmente delicato via via che ci si avvicina nel tempo: il problema dell'attribuzione di risultati a specifici ricercatori. Pur non negando l'interesse di indagini accurate al riguardo, dobbiamo confessare che in molti casi ci siamo «fidati» delle attribuzioni correnti senza compiere indagini specifiche.

Questo avrà sicuramente portato ad errori di cui ci scusiamo in anticipo, avvertendo il lettore di prendere sempre le indicazioni al riguardo *cum grano salis*. Sicuramente siamo incorsi in errori, omissioni o inesattezze in molte altre occasioni. Ringraziamo fin d'ora chi vorrà essere così gentile da segnalarceli. D'altra parte – lo vogliamo ribadire ancora una volta – questo lavoro più che un contributo di carattere storiografico o filologico vuole essere un invito ed una introduzione alla problematica della logica matematica attraverso la sua storia. Questo potrà sembrare un obiettivo modesto, anche se non possiamo negare di aver imparato molto da questa nostra fatica. Vogliamo augurarci che lo stesso valga per i nostri fatidici sette lettori.

CORRADO MANGIONE e SILVIO BOZZI

Milano, dicembre 1992

Il lavoro è frutto di una stretta collaborazione fra i due autori e quindi in generale sarebbe difficile, se non impossibile, distinguere l'apporto dell'uno o dell'altro. Per quanto però riguarda le parti nuove, vanno attribuiti esclusivamente a Silvio Bozzi i paragrafi «Dall'algebra della logica alla logica algebrica» (cap. II, § 7), «La formalizzazione della logica e della matematica intuizioniste» (cap. V, § 4), «Teoria degli insiemi da Gödel a Cohen» (cap. VI, § 6) e «La teoria degli insiemi dopo Cohen» (cap. VII, § 1).

CAPITOLO PRIMO
MUTAMENTI NEL PENSIERO MATEMATICO

1. INTRODUZIONE

Una descrizione completa e adeguata dello sfondo in cui la logica matematica è sorta alla metà del secolo scorso richiederebbe un'analisi lungo diverse direzioni tanto in ambito filosofico, quanto in ambito matematico. Il deciso allargarsi delle prospettive che la ricerca matematica registrò in questo periodo – allargamento che riguardava tanto la geometria, quanto l'algebra, l'Analisi, e la stessa matematica applicata – ebbe un ruolo centrale in questo processo ed è chiaro che in questa sede non possiamo compiere un'analisi di questo tipo. Ribadendo che solo un panorama di tutto intero lo sviluppo del pensiero matematico della prima metà dell'Ottocento potrebbe fornire l'ambientazione adeguata, ci limiteremo così a considerare più da vicino i mutamenti che riguardano il sorgere delle geometrie non euclidee e lo svilupparsi dell'algebra simbolica. Se infatti la geometria, a partire dagli *Elementi* di Euclide ha sempre costituito il prototipo della teorizzazione matematica e quindi – almeno dai greci in poi – il terreno su cui si è confrontata l'analisi logica, fu l'algebra per il suo ruolo di scienza formale indipendente dal contenuto geometrico e numerico a fornire un modello per la nozione stessa di logica come scienza delle *forme* inferenziali. Dei grandi mutamenti nel campo dell'Analisi avremo occasione di parlare più avanti, soprattutto in riferimento all'analisi del continuo e al sorgere della teoria degli insiemi e della Grande Logica che vedremo nel cap. III.

È opportuno però, fin da ora fare alcune osservazioni generali sulla matematica del primo Ottocento.

Se il grande evento nella matematica settecentesca era stato lo sviluppo tumultuoso dell'Analisi, col suo predominio sugli altri rami della matematica, nella prima metà dell'Ottocento alla otti-

mistica professione di fede dalembertiana (*allez en avant, la foi vous viendra*) che aveva fatto accumulare tanti e significativi risultati all'Analisi infinitesimale, subentrò uno spirito di attento vaglio critico di tali risultati e iniziò una profonda riflessione sui fondamenti stessi del procedere analitico. Questa esigenza si può compendiare dicendo che la ricerca matematica diveniva consapevole della necessità di una giustificazione di tipo «logico» dei suoi fondamenti, tendeva cioè ad eliminare il rimando a modelli privilegiati non concettualmente determinati (di natura geometrica o genericamente intuitiva) come sostegno delle proprie teorie. Iniziava insomma in questo periodo quel ripensamento di tutta la matematica che attraverso l'acquisizione della nozione di teoria come sistema ipotetico deduttivo avrebbe portato a quella netta separazione tra il problema dell'applicazione e quello della giustificazione delle teorie che caratterizza, nel bene e nel male, la matematica moderna.

Naturalmente l'enucleazione di tali concetti, pur avvertibile al di là di ogni dubbio in tutto l'atteggiamento matematico di questo periodo, fu faticosa e graduale, visto che per ottenerla si dovettero superare radicate convinzioni spesso soltanto di tipo pragmatico o comunque pregiudiziale, ma questa tendenza è riscontrabile in tutte le branche fondamentali della matematica, dall'algebra all'Analisi, dalla fisica matematica alla geometria. Noi prenderemo qui in esame in particolare le vicende della geometria – il presentarsi delle geometrie non euclidee – ed il sorgere della nuova algebra simbolica.

In campo algebrico questo mutato atteggiamento porta nel continente europeo a un diverso modo di affrontare classici problemi rimasti insoluti, in particolare quello della risoluzione delle equazioni algebriche generali di grado superiore al quarto, che condurrà a una delle fondamentali teorie matematiche moderne, la teoria dei gruppi. Particolarmente interessante in relazione alla nuova impostazione della logica sarà l'apporto delle scuole algebriche inglesi: in questo periodo si comprende in Inghilterra che lo «splendido isolamento» dalla matematica continentale rischiava di inaridire la ricerca analitica e si realizza con chiarezza come tale isolamento fosse in definitiva il frutto di mero sciovinismo intellettuale. Alcuni giovani scienziati inglesi intraprendono quindi un'attiva azione affinché alla tradizionale notazione flussionale

newtoniana (o «con i puntini», come allora si diceva) venga sostituita la notazione differenziale leibniziana (o «con le D») ormai diffusa in ogni altra parte d'Europa. E questa sostituzione, lungi dal costituire un puro fatto superficiale esaurentesi nel semplice rimpiazzamento di un simbolismo più agevole a un altro macchinoso e ormai superato, è in effetti all'origine, e costituisce quasi un pretesto, per la esplicazione di tutto un atteggiamento di pensiero della scuola degli algebristi di Cambridge; questi vengono progressivamente staccando i concetti algebrici dalle correnti e usuali interpretazioni numeriche o geometriche e si adoperano faticosamente, pur fra incertezze, oscurità e ripensamenti, alla introduzione di una concezione della matematica non più come teoria delle grandezze, ma in generale come teoria astratta. Se si tiene conto che proprio in questo ambiente avviene la formazione scientifica di Boole (che come vedremo è uno dei padri fondatori della logica matematica), il quale anzi dà i suoi primi contributi quale matematico proprio in questa direzione, si comprende come l'ambiente inglese di questo periodo si presenti come un nodo di centrale importanza per lo sviluppo della logica moderna.

In un analogo ordine di idee, il concetto di teoria deduttiva o formale sarà il lontano punto d'arrivo anche delle ricerche relative alle geometrie non euclidee, che come dicevamo vengono costituite proprio attorno agli anni trenta dell'Ottocento. Questa scoperta riguardava la logica non solo in quanto le poneva in modo implicito (ma diremmo anche perentorio) nuovi problemi, originati dal fatto che per questi nuovi e «strani» sistemi veniva a cadere la possibilità di una verifica «empirica» di non contraddittorietà e quindi un criterio «esterno» di correttezza del ragionamento che doveva così venir sostituito da un'analisi *logica* indipendente da riferimenti a modelli prefissati. Ma la riguardava anche e soprattutto perché comportava in modo del tutto conseguente l'apertura verso una concezione delle teorie matematiche come sistemi ipotetico-deduttivi. E come a partire dalle ricerche algebriche in Inghilterra sarà Boole a porre in modo chiaro e consapevole e da un'angolazione squisitamente logica questa nuova esigenza, così sarà lo sviluppo delle geometrie non euclidee e della geometria proiettiva che porterà David Hilbert, all'inizio del nostro secolo, a proporre una nuova visione del metodo assiomatico a fondamento della matematica moderna.

Tutto ciò, come vedremo nel prossimo capitolo, avrà grandi conseguenze sullo sviluppo della logica. Ai primi dell'Ottocento infatti assistiamo da una parte agli ultimi tentativi di sistemazione su base intensionale della logica tradizionale (ad esempio con Gian Francesco Castillon, 1708-91) e dall'altra a una decisa apertura verso il momento estensionale, sempre però nel quadro della tradizione che per comodità chiameremo settecentesca, con Gergonne nel continente e praticamente con tutti i logici inglesi «classici» (Bentham, William Hamilton, ecc.); tutti questi tentativi sfoceranno nel sistema di Boole ove troveremo (accanto ovviamente ad una trattazione estensionale) un netto rovesciamento di posizioni rispetto all'ambiente matematico: dopo che l'algebra si era in certo senso «logicizzata», è la logica che si algebrizza ponendosi, come sopra si è accennato, come nuovo ramo autonomo accanto ai classici sistemi della matematica riguardati con questa nuova e feconda mentalità.

Che così venisse colta l'intuizione leibniziana dell'importanza e della fecondità della pura ricerca formale è fuori dubbio; notevole è tuttavia il fatto – già sopra osservato – che la logica si staccava decisamente dalla filosofia, ripercorrendo del resto l'itinerario di tutte le scienze che storicamente erano venute originandosi da una comune matrice filosofica. Non mancano naturalmente in questo periodo, e specie nel continente, produzioni di logica «filosofica», tutte più o meno ispirate o comunque collegate al pensiero kantiano; esse sono anzi numerosissime al punto di costituire secondo Venn, una vera e propria «inflazione». Dopo però che la teorizzazione kantiana era stata ripresa in vari modi dalla successiva elaborazione idealista, le «logiche» di Hegel, Fichte, Schelling, ecc. ebbero come risultato sostanziale quello di bandire in modo completo e apparentemente definitivo il discorso «formale» dalla ricerca filosofica e per noi, che vogliamo qui seguire l'evoluzione della logica *formale*, esse perdono quindi ogni ragionevole interesse, così che ci limiteremo se del caso a brevissimi accenni.[1]

[1] Per la stessa ragione, anche se qui la separazione è più difficile, non ci occuperemo degli sviluppi della logica induttiva che proprio in questi anni nasce, per opera di Mill, Venn, ecc. Per questo argomento rimandiamo il lettore per esempio a D. Costantini, *Fondamenti del Calcolo delle probabilità*, Feltrinelli, 1970 e alla bibliografia ivi contenuta.

2. LA GEOMETRIA COME PARADIGMA DI TEORIA MATEMATICA

Non c'è dubbio – come ricordato sopra – che la sistemazione euclidea della geometria, con eventuali miglioramenti, abbia costituito fino almeno alla prima metà dell'Ottocento il prototipo di una scienza che da una parte offre l'esempio di un procedere deduttivo metodico, dall'altra un puntuale aggancio all'intuizione spaziale del mondo esterno. Questo malgrado fin dai tempi di Euclide (se non anche prima, come avremo occasione di ricordare più avanti) dubbi e perplessità sullo speciale tipo di rapporto fra assunzioni geometriche (postulati) e mondo esterno si fossero già presentati. L'esempio più cospicuo di queste perplessità è il modo con cui Euclide introduce il suo quinto postulato e il susseguirsi di tentativi di dimostrarlo (a partire dagli altri) o di sostituirlo con altre ipotesi che vedremo succintamente più avanti. Ciò non toglie che paradossalmente sia stato proprio col sorgere della scienza moderna che l'immagine dello spazio e della geometria sottogiacente sia andata sempre più cristallizzandosi, cosicché gli *Elementi* di Euclide divennero una sorta di «leggi divine» imposte allo spazio, la cui intuibilità, per noi, è garanzia della stessa possibilità di conoscere.

A questa concezione «intuitiva» della geometria come noto darà una perfetta sistemazione teorica Kant. Spazio e tempo sono intuizioni pure, e la geometria euclidea, in quanto articola la struttura del nostro spazio fisico quale a noi è dato concepire in forza della struttura del nostro pensiero, non può essere che quella che in effetti è: una costruzione assoluta, fondata su principi altrettanto indubitabili e assoluti. Per Kant la geometria costituiva un esempio paradigmatico di conoscenza sintetica a priori: certa in un modo che non richiede di essere giustificato dall'esperienza (ossia a priori) essa è tale che i suoi teoremi ci dicono pur qualcosa intorno al mondo, cioè non sono puramente analitici. Ma Kant, ripetiamolo, vedendo nella geometria un nesso indissolubile tra ragionamento ed intuizione, non farà che sistemare filosoficamente una concezione che si era tramandata fin dall'antichità greca per quanto riguarda la natura della geometria come interprete fedele e assoluta della struttura dello spazio fisico. Si comprende quindi come, sulla base di questa convinzione, i tentativi succedutisi nei secoli di risolvere il problema relativo al quinto postulato

degli *Elementi* fossero sostanzialmente tutti rivolti a ridurre il postulato stesso a teorema o a sostituirlo con proposizioni più evidenti, ponendo in ombra le oscillazioni che prima e dopo di Euclide c'erano state.

La geometria euclidea si presenta come studio di *figure*, punti, cerchi, triangoli, ecc. Ma dove si situano queste figure? Non dobbiamo dimenticare che la matematica greca aveva esperienza non solo della geometria del piano, ma anche di quella sferica in cui risultati validi nel piano (quali quello sulla somma degli angoli interni di un triangolo) non valgono più; aveva esperienza di angoli che non erano piani (gli angoli solidi), di linee che pur non essendo parallele non si incontravano se non asintoticamente. In più, non va dimenticato, la stessa definizione di linea retta euclidea non comporta l'infinità e solo per via di un postulato ogni linea si può prolungare indefinitamente. Quello che conta è che, come ha posto in luce I. Toth in lavori fondamentali ove esamina da questo punto di vista il *Corpus aristotelicum*, convivono nella geometria greca da una parte la considerazione delle figure in sé, a prescindere dal loro ambiente, come idee platoniche, e dall'altra la considerazione dello spazio in cui queste si situano nella nostra esperienza. La validità del quinto postulato è strettamente legata alle scelte che si compiono tra queste due concezioni, delle figure come enti in sé o delle figure come situate in uno spazio le cui proprietà sono astratte dal nostro ambiente circostante. Lo sviluppo del pensiero geometrico permetteva di vedere sempre meglio le diverse implicazioni di queste concezioni cosicché si cominciò chiaramente a comprendere che il vero problema era quello di chiedersi *se* il postulato euclideo dovesse essere necessariamente vero, *se* il postulato fosse dimostrabile, invece di ritenerlo assolutamente vero a priori o di impegnarsi testardamente nel tentativo di dimostrarlo. Questa corretta impostazione del problema, per noi oggi tanto naturale, comporta in effetti tutto un ripensamento sulla natura degli assiomi e dei postulati, sul significato stesso dell'intera geometria nei rapporti con l'esperienza e richiedeva, per essere raggiunta, un profondo mutamento di mentalità.

Una volta postesi le domande precedenti con disposizione, libera da preconcetti, ad accettarne tutte le implicite conseguenze, si giungerà appunto alla costituzione delle geometrie non euclidee, ossia si arriverà a vedere che il quinto postulato non è che una di

molte possibili ipotesi, tutte coerenti con i rimanenti postulati, che permette di edificare una teoria delle parallele; verrà così chiarito nel contempo che, lungi dal dipendere dai rimanenti postulati, una tale ipotesi è indispensabile per lo svolgimento della teoria stessa e che gli spazi in cui le figure possono situarsi non sono necessariamente univocamente determinati.

Gauss fu tra i primi a riconoscere concretamente che solo un'indagine empirica dello spazio può farci decidere della geometria che può meglio descriverlo: una volta realizzato che le geometrie non euclidee possono essere logicamente coerenti, non siamo più autorizzati a decidere, senza opportuni esperimenti empirici, quale geometria valga in natura (a questo proposito risulterà di particolare importanza e utilità il teorema sulla somma degli angoli interni di un triangolo che incontreremo spesso nelle pagine seguenti). Gauss fu dunque il primo a «pensare», se così si può dire, in modo non kantiano intorno al problema complessivo della struttura dello spazio e ciò in un periodo – la fine del Settecento – nel quale l'idea di una indagine *empirica* circa la struttura dello spazio stesso veniva considerata praticamente un assurdo.

Ci è parso quindi opportuno dare spazio fin da ora al problema dell'indipendenza del quinto postulato euclideo (e alla connessa scoperta delle geometrie non euclidee), in quanto questo problema – come ricordato – occupa, per i successivi sviluppi della matematica, della logica e della stessa filosofia, una posizione centrale che può essere così brevemente illustrata.

In primo luogo, per quanto riguarda la matematica, si realizza che anche la geometria è una scienza empirica, o quanto meno che va fatta una distinzione fra geometria «matematica» e geometria «fisica». Questo mutamento di prospettiva verrà compiutamente espresso da Hilbert nelle *Grundlagen der Geometrie* (*Fondamenti della geometria*) del 1899, dove viene codificata l'idea che era venuta progressivamente imponendosi delle costruzioni matematiche come sistemi ipotetico-deduttivi, le cui proposizioni fondamentali cioè non trascrivono alcuna verità assoluta, ma sono semplicemente ipotesi, «cominciamenti» di costruzioni formali, e non princípi indubitabili.

In secondo luogo, per quanto riguarda la logica, quella della scoperta delle geometrie non euclidee è, a nostro parere, una delle

componenti indirette della massima importanza per la sua rinascita. È infatti chiaro, in vista anche di quanto detto sopra, che se i principi di una teoria perdono il carattere di verità evidente garantita dall'intuizione, l'unica garanzia che ci resta per considerare una tale teoria come una costruzione matematica lecita è almeno la sua non contraddittorietà, ossia la certezza, di natura puramente logica, che se i principi in questione sono non contraddittori i nostri successivi passi di derivazione dei teoremi conservino per così dire questa caratteristica, siano cioè rigorosamente sottoposti al vaglio di una operante ed efficace logica deduttiva.

Infine, per quanto riguarda la filosofia, come già si può desumere da quanto fin qui detto, la scoperta delle geometrie non euclidee costringe a rivedere l'idea che il concetto di spazio sia kantianamente dato da un'intuizione pura a priori, rendendo evidente che si possono escogitare ed eseguire esperimenti per stabilire che tipo di spazio (euclideo o no, ed eventualmente con quali caratteristiche) sia lo stesso spazio fisico della nostra esperienza.

Chiarita così brevemente l'importanza del problema per la scienza e la filosofia moderne, confidiamo che il lettore sia convinto dell'opportunità, prima di prendere in esame lo sviluppo settecentesco di questo problema, di ridare uno sguardo generale d'assieme all'opera euclidea e di ricapitolare i ripetuti tentativi compiuti dai vari autori antecedenti al periodo che qui ci interessa per dimostrare il postulato delle parallele.

2.1 *Gli* Elementi *di Euclide*

Euclide non fu certamente il primo geometra greco che si accinse alla composizione di *Elementi* di geometria. Era per lo meno stato preceduto dal famoso Ippocrate (di poco anteriore a Platone, e quadratore delle celebri «lunule») e da Teudio di Magnesia, contemporaneo di Aristotele. Ma con i suoi *Elementi* Euclide riuscì a stabilire il modello tipico, destinato a rimanere tale per oltre duemila anni, del rigore dimostrativo cui era giunta la matematica greca, fornendo una ineguagliabile esemplificazione di quella che può essere chiamata l'assiomatica antica. In tale assiomatica si riconosce la convergenza di due distinti criteri, l'uno,

quello di dimostrabilità, puramente logico, l'altro, quello di evidenza, di natura extralogica, che concorrono a costituire e caratterizzare l'insieme delle proposizioni vere di una teoria. Si tratta infatti, dopo aver stabilito dei termini e delle proposizioni primitive la cui intelligibilità, rispettivamente, verità, viene garantita dall'*evidenza*, di ricavare ogni altra proposizione della geometria *dimostrandola* a partire dalle proposizioni e in base ai termini così ammessi (in funzione dei quali deve essere definito ogni altro termine introdotto nella teoria). Questa struttura logica di massima (sulla quale torneremo dopo aver illustrato brevemente il contenuto dell'opera euclidea) viene esemplificata nei 13 libri che costituiscono gli *Elementi* (un quattordicesimo e un quindicesimo libro, un tempo considerati facenti parte dell'opera, sono sicuramente spuri).

Il libro I (che ci interesserà più da vicino nel seguito) si apre con 23 definizioni di nozioni elementari quali punto, retta, superficie, angolo, rette parallele, ecc. cui segue l'elencazione dei 5 postulati e delle 5 «nozioni comuni» (assiomi) che reggeranno l'intera costruzione euclidea. Vi vengono quindi dimostrati rigorosamente 48 teoremi che costituiscono la teoria elementare dei triangoli, delle parallele e dell'equivalenza di poligoni.

Il libro II si apre con 2 definizioni e contiene 14 proposizioni di «algebra geometrica»: si tratta cioè di teoremi costituenti la versione geometrica di problemi la cui soluzione in termini algebrici comporta la considerazione di equazioni di primo e secondo grado.

Il libro III inizia con 11 definizioni di nozioni quali cerchi uguali, cerchi tangenti, arco di cerchio, corda, settore circolare, ecc. e contiene 37 proposizioni ad esse relative, costituenti la teoria elementare dei cerchi.

Il libro IV, che inizia con 7 definizioni (figure circoscritte e inscritte in un cerchio, ecc.), tratta in 16 proposizioni la teoria dei poligoni inscritti e circoscritti a una circonferenza.

Il libro V, il cui contenuto risale certamente a Eudosso, si apre con 18 definizioni e tratta, in 25 teoremi, della teoria delle proporzioni da un punto di vista generale, ossia per grandezze qualsiasi.

Alle 4 definizioni iniziali del libro VI seguono 33 proposizioni nelle quali viene sviluppata per figure geometriche partico-

lari la teoria delle proporzioni data in generale nel libro precedente.

I successivi libri VII, VIII e IX sono strettamente collegati fra loro; solo il primo di essi si apre infatti con 22 definizioni di concetti numerici (unità, numero, numero pari, numero dispari, numero primo, numero perfetto, ecc.) e i teoremi dei tre libri sono tutti relativi a questi concetti. In particolare, il libro VII tratta nelle sue 39 proposizioni della teoria della divisibilità e della teoria delle proporzioni fra interi; i libri VIII e IX, rispettivamente con 27 e 36 proposizioni, proseguono la trattazione della teoria delle proporzioni, con particolare riguardo alle «proporzioni continue» ossia alle progressioni geometriche. Vogliamo ricordare che il libro IX contiene fra gli altri alcuni famosissimi teoremi euclidei, come quello sulla unicità della scomposizione in fattori primi (proposizione 14) o quello sulla infinità dei numeri primi (proposizione 20).

Il libro X, di lettura particolarmente difficile e per la materia trattata e per il linguaggio geometrico impiegato, si apre con 4 definizioni (grandezze commensurabili e incommensurabili, razionali e irrazionali) e svolge ben 115 proposizioni della teoria degli irrazionali quadratici e biquadratici.

I rimanenti libri XI, XII e XIII formano ancora un tutto unico dedicato in generale alla geometria solida. Solo il primo di essi si apre infatti con 28 definizioni di nozioni appunto di stereogeometria, la cui teoria viene quindi svolta nelle 39 proposizioni di questo libro e nelle 36 dei rimanenti due (con 18 proposizioni ciascuno).

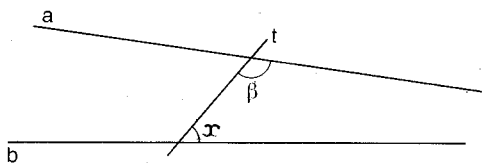
La struttura logica degli *Elementi* è così venuta implicitamente chiarendosi in questa succinta descrizione del loro contenuto: oltre ai termini che hanno funzione definitoria e che compaiono all'inizio di ogni libro, o gruppo di libri che trattano dello stesso argomento (ad esempio, del gruppo costituito dai libri VII, VIII e IX), all'inizio del solo libro I compaiono gli assiomi o nozioni comuni e i postulati.[2] Su questi assiomi e questi postulati è fondato

[2] Secondo una distinzione oggi in generale non più riconosciuta dalla logica moderna, gli assiomi sono quelle proposizioni primitive che enunciano affermazioni che si ritengono evidentemente vere in generale, indipendentemente cioè dal campo particolare cui vengono di volta in volta applicate (ad esempio «il tutto è maggiore della parte» che è il quinto assioma euclideo); i postulati invece sono proposizioni primitive la cui verità è sì assunta come evidente, ma la cui validità

l'aspetto più propriamente deduttivo degli *Elementi*, nel senso che ogni altra proposizione dell'opera viene accettata in forza di una rigorosa dimostrazione basata su quelle proposizioni iniziali (o, ovviamente, su proposizioni già dimostrate); in altri termini, la verità delle ulteriori proposizioni geometriche viene ricondotta dimostrativamente alla verità delle proposizioni iniziali assunta come evidente, e viene da questa garantita.

I cinque postulati che Euclide pone alla base della sua costruzione sono i seguenti:

«Si postula che:

1) da qualsiasi punto si possa condurre una retta a ogni altro punto;
2) ogni retta terminata [ossia ogni segmento] si possa prolungare continuamente, per diritto;
3) con ogni centro e ogni distanza si possa descrivere una circonferenza;
4) tutti gli angoli retti sono uguali fra loro;
5) se una retta, incontrandone altre due, forma gli angoli interni da una stessa parte minori di due retti, le due rette prolungate all'infinito si incontrino, dalla parte in cui sono i due angoli minori di due retti.»

Con riferimento alla situazione illustrata in figura, il postulato 5 richiede che se $\alpha + \beta < 180°$, allora le due rette *a* e *b* si incontrino dalla parte di α e β.

Non affrontando qui il problema di altre deficienze che la critica moderna ha messo in luce relativamente alla costruzione assiomatica euclidea, il «neo» per eccellenza fu considerato dagli stessi primi commentatori di Euclide appunto quello dell'ammissione dell'ultima delle proposizioni precedenti come postulato: essa non

è per così dire limitata al campo specifico della scienza che li assume e che consistono generalmente nell'ammettere la possibilità di effettuare determinate costruzioni.

ha infatti le stesse peculiari caratteristiche di «evidenza» godute dalle altre quattro; è anzi estremamente probabile che tale fosse anche la convinzione di Euclide, il quale non impiega il postulato in questione se non nella proposizione 29 del libro I (in linguaggio moderno: due rette parallele tagliate da una trasversale formano angoli alterni uguali, angoli corrispondenti uguali, angoli coniugati supplementari) dopo di che esso diventa anche operativamente una delle basi fondamentali per l'intero sistema. Va notato che nella definizione 23 del libro I Euclide aveva definito «parallele» due rette che prolungate indefinitamente in entrambe le direzioni non si incontrano mai; si osservi infine che il quinto postulato viene spesso presentato (oltre che come «postulato delle parallele» o semplicemente «postulato di Euclide») come «postulato dell'unicità della parallela» in quanto la proposizione che, in un piano, a una retta data possa condursi una, e una sola, parallela per un punto fuori di essa, si ricava dalla proposizione 30 del libro I (rette parallele a una stessa retta sono parallele fra loro) e può dimostrarsi logicamente equivalente al quinto postulato.

Questa equivalenza può aiutare a meglio comprendere il difetto di evidenza lamentato nella proposizione euclidea: è chiaro infatti che in una regione piana, per quanto grande ma comunque accessibile ai nostri sensi o alla nostra intuizione diretta, dati una retta e un punto fuori di essa noi possiamo realmente pensare a infinite rette che passano per quel punto e *non* incontrano la retta data (discorso opposto se pensiamo di muoverci su di una superficie sferica, ossia la superficie che fornisce il modello del cosmo greco); ora, affermare che al di fuori del campo della nostra possibile esperienza (campo che, per quanto grande, è pur sempre limitato) tutte quelle rette *meno una* incontrino la retta data (o che nel secondo esempio tutte s'incontrino) è indubbiamente una impegnativa estrapolazione che ci porta ben al di là dei dati dell'osservazione e dell'intuizione.

2.2 *La critica al quinto postulato fino a Saccheri*

Come si è poco sopra accennato già i primi commentatori degli *Elementi* cercarono di eliminare dalla sistemazione euclidea il «neo» rappresentato dall'assunzione del quinto postulato (costi-

tuendo una relativa teoria rigorosa delle parallele) dando così origine a tutta una serie di tentativi in questo senso che, succedutisi nel corso dei secoli, portarono infine alla scoperta delle geometrie non euclidee nei primi decenni del XIX secolo. Per quanto talora estremamente diversi, tutti questi vari tentativi possono sostanzialmente farsi rientrare in uno dei seguenti tre tipi, non necessariamente escludentisi fra loro: 1) assunzione di una definizione di rette parallele diversa da quella euclidea; 2) sostituzione del quinto postulato con un'altra proposizione più intuitiva, ossia la cui verità risultasse più «evidente», e quindi di più facile accettazione; 3) dimostrazione del postulato come teorema, deducendolo dai quattro postulati rimanenti. In questo paragrafo accenneremo brevemente ai più importanti tentativi nelle varie direzioni che si possono riscontrare fino al XVIII secolo.

Per quanto riguarda l'antichità, la nostra migliore fonte d'informazione in proposito è Proclo (410-485) che prima di presentare una propria personale proposta di soluzione riferisce su vari tentativi di predecessori, a cominciare da quello di Posidonio (I secolo a.C.). Questi tenta di superare il problema dando una diversa definizione di parallelismo: due rette sono parallele quando sono complanari ed equidistanti. Con un tale artificio però non si fa che spostare la questione, non la si risolve: in tal caso infatti si rende necessario o *dimostrare* che il luogo dei punti equidistanti da una retta e giacenti da una parte di essa è a sua volta una retta (e questa dimostrazione richiede l'impiego del postulato euclideo) oppure si deve *postulare* l'esistenza di coppie di rette che godano di queste proprietà e quindi assumere un nuovo postulato certamente non più evidente di quello euclideo. Altro tentativo riferito da Proclo è quello di Tolomeo (II secolo d.C.) il quale tenta di dimostrare il quinto postulato deducendolo dalla proposizione che la somma degli angoli interni di un triangolo è uguale a due retti. Da questa proposizione il postulato euclideo si ottiene in effetti facilmente, ma Tolomeo, per dimostrare la proposizione da cui prende le mosse, fa l'ipotesi che se quel caso si verifica per un triangolo lo stesso si verifica per tutti, la quale ipotesi è per lo meno così poco evidente quanto lo stesso postulato euclideo.

Nel prospettare la propria proposta di soluzione del problema, Proclo dichiara che si rifiuta di assumere come postulato una proposizione (la proposizione euclidea, appunto) la cui inversa è un

teorema. Egli osserva infatti che la proposizione inversa del postulato euclideo è nient'altro che la diciassettesima del libro I degli *Elementi* la quale afferma che la somma di due angoli interni di un triangolo è minore di due angoli retti. Proclo assume quindi come nuovo postulato il fatto che la distanza fra due punti presi su rette intersecantesi può essere resa grande a piacere prolungando sufficientemente le due rette, donde deduce il lemma secondo il quale una retta che incontri una di due rette parallele deve incontrare anche l'altra. In questa dimostrazione tuttavia Proclo introduce l'ipotesi che la distanza fra le due parallele rimane finita; e da questa ipotesi si può dedurre logicamente il postulato euclideo (senza contare il fatto che già l'assunzione del nuovo postulato, che Proclo ricava da un passo di Aristotele, non risulta di certo più «evidente» di quello euclideo).

Altre critiche e vari tentativi di emendare gli *Elementi* vennero successivamente avanzati dai commentatori arabi, che introdussero ipotesi o adottarono procedimenti dimostrativi talora geniali. Qui ci limitiamo a ricordare l'opera di Anarizio (Al Narizi, IX secolo) tradotta in latino da Gherardo da Cremona (XII secolo) *Euclidis Elementa ex interpretatione Al Hadschdschadschii cum commentariis Al Narizii* (*Gli Elementi di Euclide nell'interpretazione di Al Hadschdschadsch commentati da Anarizio*) che segue sostanzialmente il procedimento di Posidonio, fondando la sua presunta dimostrazione del quinto postulato sull'assunzione che esistono due rette equidistanti, e l'originale tentativo di Nasir ed Din (1201-74) il quale si fonda su un'ipotesi (a sua volta ben poco «evidente») dalla quale è possibile derivare immediatamente che la somma degli angoli interni di un triangolo è uguale a due retti e, da questo risultato, ottenere subito il postulato euclideo (che gli è equivalente).

Durante il Rinascimento e il XVII secolo, le critiche al postulato euclideo cominciarono soltanto dopo che fu possibile disporre della versione latina dei commentari di Proclo (che vennero stampati a Basilea nel 1533 in originale e quindi in latino a Padova nel 1560) e comunque praticamente nella seconda metà del secolo. Di questo periodo ricorderemo per lo più solo dei nomi, soffermandoci brevemente su quegli autori che appaiono più significativi per eventuali spunti originali nelle loro critiche. Così F. Commandino (1509-75) negli *Elementorum libri XV* (*I quindici libri degli Elementi*) stampati a Pesaro nel 1572 segue la dimostrazione di Proclo per il

quinto postulato, dopo aver aggiunto l'idea di equidistanza nella definizione euclidea di parallele; R.S. Clavio (1537-1612) negli *Euclidis elementorum libri* XV (*I quindici libri degli Elementi di Euclide*) stampati a Roma nel 1574, si rifà a sua volta a Proclo criticandolo, quindi tenta di dare una propria dimostrazione del postulato euclideo fondata sull'ipotesi che la linea equidistante da una retta è a sua volta una retta seguendo sostanzialmente la dimostrazione di Nasir ed Din. La prima opera esclusivamente dedicata al problema delle parallele è l'*Operetta delle linee rette equidistanti et non equidistanti* (Bologna 1603) di P.A. Cataldi (1552-1626) ove però l'autore, per la dimostrazione del quinto postulato, si fonda su un'ipotesi già avanzata da Nasir ed Din. Giovanni Alfonso Borelli (1608-79) fa intervenire il concetto di movimento nella sua trattazione del problema e definisce le parallele come rette equidistanti nel celebre *Euclide restitutus* che pubblica a Pisa nel 1658. Degna di nota è l'opera di Giordano Vitale (1633-1711) che nel suo *Euclide restituto overo gli antichi elementi geometrici restaurati e facilitati. Libri* XV stampato a Roma nel 1680, ritorna all'idea di Posidonio riconoscendo però la necessità di escludere che rette parallele in senso euclideo siano asintotiche (ossia si avvicinino indefinitamente senza tuttavia incontrarsi); tenta quindi di dimostrare che il luogo dei punti equidistanti da una retta è ancora una retta. Nell'opera di Vitale si possono riscontrare notevoli spunti originali, alcuni dei quali verranno ripresi da Saccheri.

Uno dei più notevoli tentativi compiuti nel XVII secolo per dimostrare il quinto postulato euclideo è quello di John Wallis (1616-1703) il quale, considerata l'infruttuosità della via battuta da molti suoi predecessori fondata sul concetto di equidistanza, pone alla base della propria dimostrazione del postulato (e quindi della ricostruzione della teoria delle parallele) l'assioma secondo il quale, data una figura, ne esiste un'altra ad essa simile e di dimensioni arbitrarie. In effetti Wallis si serve di questa postulazione solo per il caso dei triangoli e Saccheri dimostrerà che assumere questo postulato è equivalente ad assumere l'ipotesi dell'angolo retto, ossia il postulato euclideo; senza contare d'altra parte che il concetto di «forma», cui implicitamente rimanda in modo immediato la nozione di «similarità» che compare nel postulato di Wallis, è fra i più complessi di tutta la geometria e abbisognerebbe comunque a sua volta di una chiarificazione preliminare.

2.3 *I «precursori» delle geometrie non euclidee*

Siamo così giunti al XVIII secolo, ossia al secolo dei veri e propri precursori delle geometrie non euclidee. Si è già accennato infatti che i tentativi secolari di dimostrazione del quinto postulato di Euclide sfoceranno nei primi decenni del XIX secolo nella costituzione di geometrie che assumono fra i loro postulati la *negazione* del postulato euclideo e che cionondimeno risultano logicamente coerenti.

I nomi che qui incontreremo saranno, oltre a quelli di Saccheri e Lambert, quelli di numerosi altri studiosi – in particolare di Adrien Marie Legendre – che, pur essendo fioriti già nei primi decenni del XIX secolo, possono tuttavia senza forzature essere associati ai precedenti per quanto riguarda la convinzione della verità assoluta del postulato euclideo, caratteristica appunto di tutta la schiera di precursori, anche in senso stretto, delle geometrie non euclidee.

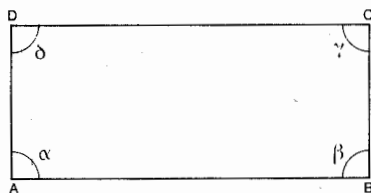
Gerolamo Saccheri (1667-1733) entra in modo del tutto originale nella preistoria delle geometrie non euclidee col suo *Euclides ab omni naevo vindicatus: sive conatus geometricus quo stabiliuntur prima ipsa universae Geometria Principia* (*Euclide emendato da ogni neo, ossia tentativo geometrico col quale si stabiliscono gli stessi primi principi di geometria universale*) stampato a Milano nel 1733, anno stesso della morte del suo autore. Oltre al neo fondamentale dal quale Saccheri vuole emendare l'opera di Euclide, la questione del postulato delle parallele, egli ne prospetta un altro, relativo alla teoria delle proporzioni (del libro V degli *Elementi*) di cui però non ci occuperemo qui dato che Saccheri non dimostra, nella trattazione di questo secondo problema, un'originalità e un rigore paragonabili a quelli esibiti nella trattazione del primo. I motivi dell'originalità di Saccheri rispetto ai suoi predecessori (e sostanzialmente anche rispetto ai contemporanei e successori) sono essenzialmente due:

1) egli affronta la questione da logico, in relazione alle ricerche della *Logica demonstrativa* del 1697, ossia non cerca di sostituire il postulato con una diversa ipotesi o di variare opportunamente la definizione di rette parallele, ma applica il metodo di ragionamento già elaborato nella sua precedente opera per dimostrare che il quinto postulato è una conseguenza logica dei rimanenti quattro. Si tratta di negare il postulato delle parallele, e di ricavare da que-

sta negazione tutte le conseguenze logiche necessarie fino a incontrarne una (sulla cui esistenza il Saccheri non nutre alcun dubbio) che risulti incompatibile col sistema dei primi quattro postulati di Euclide. In questo modo si sarebbe dimostrato che conseguenza degli assiomi e della negazione del quinto postulato è un assurdo, quindi che dagli assiomi è ricavabile la negazione della negazione del quinto postulato cioè il postulato stesso.

2) Proprio per questo suo tipo di approccio generale, Saccheri è il primo a prendere in esame tutte le possibilità che una negazione del postulato in questione comporta e in tal modo egli può considerarsi come precursore di entrambe le geometrie non euclidee successivamente scoperte, ossia tanto della geometria ellittica quanto di quella iperbolica.

Secondo Thomas Heath, Saccheri «fu vittima della preconcetta convinzione del suo tempo che la sola geometria possibile fosse quella euclidea, ed egli presenta lo spettacolo curioso di un uomo che erige a fatica una struttura su nuove fondamenta con l'effettivo proposito di demolirla in seguito; egli cercò contraddizioni nel cuore del sistema che egli stesso aveva costruito, per dimostrare così la falsità delle sue ipotesi». Saccheri infatti assume come dati (oltre ai primi quattro postulati euclidei) le prime ventotto proposizioni del libro I (che sono indipendenti dal quinto postulato) e fa inoltre l'ipotesi che il postulato delle parallele sia falso: giusto il metodo di ragionamento da lui adottato, cerca fra le conseguenze di queste ipotesi una proposizione che lo autorizzi ad affermare la verità del quinto postulato stesso.

Saccheri prende le mosse dalla considerazione di una figura fondamentale, il quadrilatero birettangolo isoscele

ottenuto innalzando dagli estremi A e B della base AB due lati AD e BC uguali fra loro e perpendicolari alla base, e dimostra innanzitutto che i due angoli in C e in D sono uguali ($\gamma = \delta$). Nel caso

che questi angoli siano a loro volta angoli retti, siamo nell'ipotesi euclidea, sicché se assumiamo che essi siano o entrambi acuti o entrambi ottusi neghiamo implicitamente, in entrambi i casi, il quinto postulato di Euclide. Saccheri prende quindi in esame le tre seguenti ipotesi relative agli angoli γ e δ:

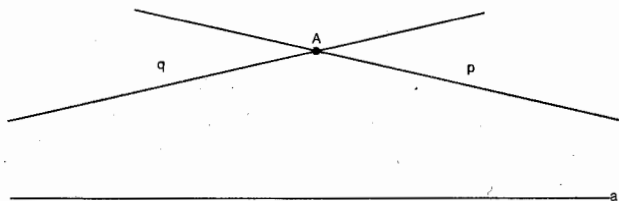
1) ipotesi dell'angolo retto: $\gamma = \delta = 90°$
2) ipotesi dell'angolo ottuso: $\gamma = \delta > 90°$
3) ipotesi dell'angolo acuto: $\gamma = \delta < 90°$.

Fra le prime interessanti proposizioni che ricava dalla sua analisi è quella secondo cui, se si può dimostrare vera in un caso singolo rispettivamente l'ipotesi 1) o 2) o 3), allora quella ipotesi vale in ogni caso (proposizioni V, VI, VII). Va quindi citata la proposizione IX, secondo la quale, a seconda che si assuma l'ipotesi dell'angolo retto o dell'angolo ottuso o dell'angolo acuto, la somma degli angoli interni di un triangolo risulta rispettivamente uguale, maggiore, minore di due angoli retti.

Un primo risultato, riguardante l'insostenibilità dell'ipotesi dell'angolo ottuso, viene ottenuto da Saccheri nella proposizione XIV. Egli aveva infatti dimostrato (proposizione XIII) che il quinto postulato è vero nel caso dell'ipotesi dell'angolo retto e in quella dell'angolo ottuso; e allora basta considerare che valgono anche tutti i teoremi deducibili da quel postulato, in particolare quello affermante che la somma degli angoli del quadrilatero fondamentale è di quattro angoli retti; viene quindi a cadere l'ipotesi 2) che richiederebbe ovviamente che questa somma fosse maggiore di quattro retti, donde si ricava che è vera l'ipotesi dell'angolo retto; in altri termini, afferma Saccheri, l'ipotesi dell'angolo ottuso «distrugge se stessa». In effetti il procedimento di Saccheri (e quindi la sua stessa conclusione) non è corretto. La sua dimostrazione infatti prova semplicemente che l'ipotesi dell'angolo ottuso non è compatibile col sistema complessivo delle altre premesse della geometria che egli ha assunto e richiede quindi che almeno una di esse – diversa naturalmente dal quinto postulato – venga abbandonata. In particolare Saccheri nella dimostrazione della citata proposizione XIII fa intervenire le proposizioni 16 e 18 del libro I degli *Elementi*, che valgono solo sotto l'assunzione che la retta sia infinita, e tale assunzione appunto non è valida sotto l'ipotesi dell'angolo ottuso.

Comunque sia, conclusa la confutazione della seconda ipotesi senza difficoltà particolari, si tratta ora per Saccheri di stabilire la falsità della terza ipotesi, quella dell'angolo acuto. Il problema si presenta qui subito più complesso, tanto che egli parla di «diuturnum proelium contra hypotesim anguli acuti»; e sarà proprio in occasione di questa dimostrazione che Saccheri verrà meno a quel sottile rigore logico che aveva sin qui contraddistinto la sua opera. Senza seguire dettagliatamente i vari passi della complessa confutazione della terza ipotesi, ci limitiamo a riassumerli come segue.

Nelle proposizioni XXIII e XXV Saccheri dimostra che, nella ipotesi dell'angolo acuto, due rette complanari o si incontrano, o ammettono una perpendicolare comune o infine, da una determinata banda, esse vanno sempre più avvicinandosi l'una all'altra in modo che la loro distanza diventa minore di qualsiasi segmento piccolo a piacere. In altri termini, in quest'ultimo caso le due rette si comportano asintoticamente.

Stabilita l'esistenza di rette asintotiche (si potrebbe dire di «parallele non euclidee») nelle proposizioni XXX, XXXI e XXXII giunge a risultati che complessivamente possono essere presentati come segue: nell'ipotesi dell'angolo acuto, nel fascio di rette per un punto esistono due rette asintotiche ad una retta data che dividono le rette del fascio in due classi; alla prima classe appartengono quelle rette del fascio che incontrano la retta data, alla seconda quelle che hanno con essa una perpendicolare comune. Questa proposizione è illustrata dalla figura seguente, dove A è il centro del fascio, a la retta data e p e q sono le due rette asintotiche.

Stabiliti questi teoremi (che verranno «riscoperti», ovviamente in modo indipendente, dai fondatori della geometria iperbolica) Saccheri ha urgenza di concludere: il suo ragionamento, finora assai rigoroso, slitta su un piano intuitivo e di puro convincimento psicologico quando trasferisce all'infinito concetti e considerazioni

validi per enti posti a distanza finita (cadendo nei facili equivoci che questo trasferimento comporta) per dimostrare, nella proposizione XXXIII, che l'ipotesi dell'angolo acuto è assolutamente falsa «perché ripugna alla natura della retta». Le considerazioni a questo scopo sono contorte e spesso oscure, ma nel complesso equivalgono all'affermazione che se l'ipotesi dell'angolo acuto fosse vera allora le rette *a* e *p* della figura precedente avrebbero una perpendicolare comune nel loro punto comune all'infinito, il che appunto – a suo parere – è contrario alla natura della retta.

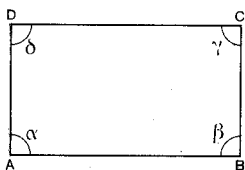
Saccheri stesso non fu soddisfatto di questa conclusione, tanto è vero che cercò di dimostrare di nuovo l'insostenibilità della terza ipotesi ricorrendo a una argomentazione fondata sul vecchio concetto di equidistanza; ma in questo suo secondo approccio non ottenne alcun risultato che possa qui interessarci e che comunque aggiunga qualcosa ai meriti, peraltro grandissimi, di Saccheri.

L'opera di Saccheri suscitò indubbiamente larga eco fra i suoi contemporanei, anche se cadde ben presto nella dimenticanza più assoluta. Essa ad esempio viene citata nella storia della matematica di J.C. Heilbronner (1706-45), edita a Lipsia nel 1742, come pure nella notissima storia della matematica di Jean-Étienne Montucla (1725-99), pubblicata a Parigi nel 1758. Va menzionata però in particolare la citazione che ne fa Georg S. Klügel (1739-1812) nel suo esame di una trentina di «dimostrazioni» del quinto postulato euclideo.

Il lavoro di Klügel è rimarchevole soprattutto per la sua conclusione che sembra per la prima volta avanzare qualche dubbio sulla dimostrabilità del quinto postulato, sostenendo che la certezza della sua verità è più frutto di osservazioni sperimentali che non di rigorosa dimostrazione; si tratta del noto *Conatum praecipuorum theoriam parallelarum demonstrandi recensio, quam publico examini submittent A. G. Kaestner et auctor respondens G. S. Klügel* (*Rassegna dei principali tentativi di dimostrare la teoria delle parallele, che viene sottoposta al pubblico esame da A. G. Kaestner e dall'autore responsabile G. S. Klügel*), pubblicato a Gottinga nel 1763 e molto citato da Johann Heinrich Lambert (1728-77) nella sua *Theorie der Parallelinien* (*Teoria delle parallele*) composta nel 1766 ma pubblicata postuma solo nel 1782 da Bernoulli e Hindenburg. Si può dunque senz'altro affermare che Lambert conoscesse l'opera di Saccheri e la cosa sarebbe anche confermata – come vedremo – dall'analogia della figura fonda-

mentale che Lambert stesso pone alla base della sua indagine. Resta comunque il fatto che Lambert riuscì ad andare molto più al di là di Saccheri nello sviluppare le conseguenze delle due ultime ipotesi; probabilmente il non aver pubblicato le proprie riflessioni sull'argomento sta a significare per l'appunto che Lambert nutriva un ragionevole dubbio circa la validità incondizionata del postulato euclideo: ma tanto su questa ipotesi quanto sulla precisazione di una eventuale effettiva influenza dell'opera di Saccheri su Lambert siamo costretti ancora oggi a rimanere nell'ambito delle congetture.

L'opera sopra citata di Lambert è composta di tre parti, nella prima delle quali viene discussa la possibilità che il quinto postulato sia dimostrabile dai primi quattro o se invece occorra assumere, allo scopo, qualche altra ipotesi; nella seconda vengono presi in esame vari tentativi di ridurre il postulato in questione ad altre proposizioni semplici che richiedono tuttavia a loro volta di essere dimostrate; la terza infine, che è quella sulla quale riferiremo brevemente, contiene una ricerca analoga a quella condotta da Saccheri.

Lambert assume come figura fondamentale il quadrilatero trirettangolo, ossia con tre angoli retti, e le analoghe ipotesi di Saccheri vengono fatte sulla natura del quarto angolo.

Le tre ipotesi di Lambert (che noi riferiremo anche, come al solito, quale prima, seconda, terza ipotesi rispettivamente) sono:

1) ipotesi dell'angolo retto: $\gamma = 90°$
2) ipotesi dell'angolo ottuso: $\gamma > 90°$
3) ipotesi dell'angolo acuto: $\gamma < 90°$.

Dimostra anch'egli la validità generale di ognuna di queste ipotesi qualora sia possibile dimostrarla per un caso, e procede quindi a escludere l'ipotesi dell'angolo ottuso (è chiaro che la prima

ipotesi dà luogo al sistema euclideo) dimostrando facilmente come a partire da essa possa derivarsi una contraddizione. Per quanto riguarda la terza ipotesi, egli trova intanto che in questo caso la somma degli angoli interni di un triangolo è minore di due retti e generalizzando il ragionamento di Saccheri introduce il concetto di *difetto* di un poligono, che è dato dalla differenza fra $2(n - 2)$ angoli retti e la somma degli angoli interni del poligono stesso, dove n è il numero dei suoi lati. Dimostra quindi che questo difetto è proporzionale all'area del poligono.

Altra considerazione notevole di Lambert connessa a questo risultato, e sulla quale si baserà la refutazione dell'ipotesi dell'angolo acuto, consiste nella scoperta che in una geometria fondata su questa ipotesi è possibile assegnare una unità *assoluta* di lunghezza cosicché ogni misura di lunghezza verrebbe appunto ad avere un valore assoluto a differenza di quanto avviene nell'ordinaria geometria dove, come noto, la misura ad esempio di un segmento è relativa all'unità di misura scelta. Lambert scopre che a ogni segmento può essere associato un angolo definito e facilmente costruibile (si ricordi che per un angolo è facile dare una misura assoluta, basta esprimerlo per esempio in radianti). Ad esempio, dato un segmento, si può pensare di costruire su di esso un triangolo equilatero e quindi di assegnare al segmento l'angolo di tale triangolo (si ricordi che siamo nell'ipotesi dell'angolo acuto e che sostanzialmente affermare l'esistenza di una unità di misura assoluta significa che non possono esistere figure simili e non uguali). Si ottiene così una corrispondenza biunivoca fra angoli e segmenti che può condursi ad avere tutte le usuali proprietà della ordinaria misura, anche se a prezzo di certe, del resto elementari, complicazioni.

In definitiva, la possibilità di costruire l'unità assoluta di misura dipende dalla possibilità di costruire (sempre nell'ipotesi dell'angolo acuto!) un triangolo equilatero di dato difetto. D'altra parte possiamo generalizzare, assegnando una misura assoluta anche all'area di un poligono (semplicemente: il suo difetto) come pure al volume di un poliedro; ma, osserva Lambert, tutto ciò non si accorda con l'ordinaria intuizione dello spazio, sicché va rigettata la possibilità di stabilire misure assolute per lunghezze, aree e volumi: ne viene di conseguenza che va refutata l'ipotesi dell'angolo acuto di cui questa possibilità è conseguenza (tralasciamo

qui di prendere in esame il problema se analoghe considerazioni non possano farsi anche sotto l'ipotesi dell'angolo ottuso).

Lambert mette per primo in evidenza la stretta analogia fra geometria su una sfera e geometria piana sotto l'ipotesi dell'angolo ottuso e osserva che la geometria sferica è indipendente dal postulato delle parallele. Citiamo ancora una sua sottile, si potrebbe dire «profetica», osservazione che l'ipotesi dell'angolo acuto dovrebbe essere verificata nel caso di una «sfera immaginaria». Un suggerimento per questa sua osservazione può essergli venuto dal momento più propriamente «tecnico» delle sue riflessioni: egli dà infatti, per l'area Δ di un triangolo piano nell'ipotesi dell'angolo acuto, la seguente formula

$$\Delta = k\,(\pi - A - B - C) \tag{1}$$

e in quella dell'angolo ottuso

$$\Delta = k\,(A + B + C - \pi) \tag{2}$$

dove k è una costante. Se nella (2) si pone $k = r^2$ si ottiene l'area del triangolo sferico (su una sfera di raggio r) e se ora si opera la sostituzione $r = r\sqrt{-1}$ (ossia se ci si pone su una sfera di raggio immaginario) e si pone quindi $r^2 = k$, si ottiene proprio la (1).

Come già abbiamo accennato, ricorderemo ora alcuni autori che, sebbene da un punto di vista cronologico non appartengano tutti, a rigore, al periodo da noi preso in esame, sono tuttavia «settecenteschi» per quanto riguarda il loro approccio al problema qui esaminato. Va detto intanto che nel XVIII secolo il discorso sui fondamenti della geometria, oltre che in Italia dove Saccheri appare peraltro come isolato, si sviluppa essenzialmente in Germania, in particolare verso la fine del secolo e segnatamente sotto l'influenza di Gauss, che incontreremo fra i fondatori delle geometrie non euclidee. Ma il problema era sentito in tutta Europa. In Francia ad esempio, oltre a Legendre al quale va dedicato un discorso a parte, altri numerosi e noti scienziati non mancano di pronunciarsi sull'annosa questione. Così ad esempio Jean-Baptiste Le Rond d'Alembert (1717-83) il quale riteneva che «la definizione e le proprietà della retta così come delle parallele sono lo scoglio e per così dire lo scandalo degli elementi della geometria».

Egli propone di definire la parallela a una retta come l'altra retta ad essa complanare che unisca due punti dalla stessa parte e alla stessa distanza della retta data, e lascia in certo senso come compito ai contemporanei di dimostrare che le rette così ottenute sono equidistanti. Giuseppe Luigi Lagrange (1736-1813) scrisse addirittura una memoria sulla teoria delle parallele, che però ritirò al momento di darne lettura in una seduta dell'accademia delle scienze; Lazare N.M. Carnot (1753-1823) e Pierre-Simon de Laplace (1749-1827) studiano il problema partendo entrambi da nozioni di similarità analoghe a quelle di Wallis che vengono tuttavia giustificate, in particolare da Laplace, con considerazioni di carattere analogico meccanico. Notevole ancora la posizione di Jean-Baptiste Joseph Fourier (1768-1830) per una sua discussione sull'argomento avuta con Gaspard Monge (1746-1818), nella quale sembra assumere un atteggiamento che sarà successivamente peculiare dei fondatori delle geometrie non euclidee.

Mentre, verso la fine del secolo, gli studi per la costituzione di una rigorosa teoria «emendata» delle parallele erano orientati in Germania verso ricerche che almeno in via ipotetica si concretassero in risultati geometrici ottenuti sotto assunzioni diverse da quella euclidea, ossia in senso che preludeva direttamente la scoperta delle geometrie non euclidee, in Francia (come del resto in Inghilterra), malgrado come si è visto l'argomento fosse di piena attualità, l'indirizzo delle ricerche era in generale di tipo tradizionale: i vari autori erano cioè impegnati nel tentativo di *dimostrare* il quinto postulato. In questo senso si muovono anche le ricerche del più noto studioso francese della questione, cui abbiamo avuto occasione di accennare più volte: Adrien Marie Legendre (1752-1833). Questi compì una serie di indagini che venne via via esponendo nelle 12 successive edizioni, dal 1794 al 1823, dei suoi fortunati *Éléments de géométrie* (*Elementi di geometria*) e raccolse inoltre i contributi specifici al problema del quinto postulato in una memoria presentata nel 1833, all'accademia delle scienze di Parigi, dal titolo *Refléxions sur différentes manières de démontrer la théorie des parallèles ou le théorème sur la somme des trois angles du triangle* (*Riflessioni su differenti maniere di dimostrare la teoria delle parallele o il teorema sulla somma dei tre angoli del triangolo*).

Suo punto fermo è dunque l'idea della riduzione del postulato delle parallele a teorema, ossia la dimostrazione del quinto postu-

lato. Le sue diverse dimostrazioni si differenziano, oltre che naturalmente a livello puramente «tecnico», per le ipotesi diverse che, come vedremo, egli assume di volta in volta; esse tuttavia hanno tutte in comune il punto di partenza in quanto Legendre vuol giungere a provare il postulato ricavandolo dalla proposizione, che quindi deve preliminarmente dimostrare, che la somma degli angoli interni di un triangolo è uguale a due angoli retti.

Egli intanto comincia col refutare l'ipotesi dell'angolo ottuso dimostrando (ovviamente senza ricorrere al quinto postulato) che la somma degli angoli interni di un triangolo è non maggiore (ossia minore o uguale) a due angoli retti. Sappiamo che questa proposizione era già stata dimostrata da Saccheri nella propria refutazione dell'ipotesi dell'angolo ottuso, ma è passata nella letteratura come primo teorema di Legendre; del resto è nota nella letteratura come secondo teorema di Legendre un'altra proposizione che sappiamo già dimostrata da Saccheri, e che potremmo dire della generalizzazione, in quanto afferma che se uno dei due casi (somma minore o uguale a 180°) si verifica per un solo triangolo, allora lo stesso caso si verifica per tutti i triangoli.

Dimostrati questi teoremi, Legendre refuta l'ipotesi dell'angolo acuto, vale a dire dimostra che la somma degli angoli interni di un triangolo è proprio uguale a due retti, con un'argomentazione affine a quella usata per lo stesso scopo da Lambert, basata cioè sull'impossibilità dell'esistenza di un'unità assoluta di lunghezza. Stabilita quindi la validità dell'ipotesi dell'angolo retto, Legendre deriva il postulato delle parallele nella forma dell'unicità, ossia dimostra la seguente proposizione: se la somma dei tre angoli interni di un triangolo è uguale a due retti, per ogni punto di un piano può condursi una e una sola parallela a una retta data.

Alcune delle ipotesi cui Legendre fa di volta in volta ricorso nelle sue varie dimostrazioni sono le seguenti:

1) Da un punto preso arbitrariamente nella porzione di piano determinata da un angolo è possibile tracciare una retta che taglia entrambi i lati dell'angolo;

2) La scelta dell'unità di lunghezza non influenza la verità di ciò che viene dimostrato (analoga all'ipotesi di Wallis: esistono figure simili di dimensione arbitraria);

3) Dati tre punti non collineari esiste sempre un cerchio che passa per essi.

Come si vede dunque Legendre per «dimostrare» il postulato euclideo è costretto a far intervenire assunzioni che o sono equivalenti alla proposizione di Euclide o non ne sono certamente più «evidenti». È indubbio che Legendre abbia svolto un'importantissima funzione e per aver diffuso l'argomento, anche a livello scolastico, e per aver ottenuto tutta una serie di risultati indipendenti dalla teoria euclidea delle parallele; va anche detto però che egli non ha contribuito in senso effettivo alla successiva scoperta delle geometrie non euclidee.

Eppure verso la fine del XVIII secolo e l'inizio del XIX, questa scoperta è, per così dire, nell'aria. Ricordiamo per finire due autori, l'uno ungherese, l'altro tedesco, entrambi in contatto con Gauss, che sia pure in maniera diversa possono confermarci questa impressione. Il primo di essi è Wolfgang Bolyai (1775-1856) il quale studia a Gottinga nel periodo 1796-99 e vi è compagno di Gauss. Non ci è noto con sicurezza quando egli inizi a interessarsi al problema delle parallele; sta di fatto che tale problema lo affanna per lungo tempo e che egli giunge a intravedere, pur se vagamente, la possibilità che il postulato euclideo sia falso. Nel 1804 fa pervenire a Gauss una sua *Theoria parallelarum* (*Teoria delle parallele*) nella quale ritiene di aver dimostrato l'esistenza di rette equidistanti. Gauss scopre l'errore nella dimostrazione, ma Bolyai prosegue testardamente nei suoi studi giungendo però solo a sostituire la proposizione euclidea con altre più o meno evidenti (e che naturalmente richiedono a loro volta di essere dimostrate); una delle ipotesi più interessanti dalla quale egli vuol derivare il quinto postulato è espressa dalla proposizione: quattro punti non complanari giacciono sempre su una sfera, che equivale alla terza ipotesi su ricordata di Legendre che per tre punti non collineari passi sempre un cerchio.

Tanto accanita perserveranza di Wolfgang Bolyai verrà in certo senso compensata dal fatto che toccherà a suo figlio Janos legare il proprio nome alla scoperta della geometria non euclidea (iperbolica). Quando, nel 1823, questi comunica al padre la propria scoperta, Wolfgang lo esorta a pubblicare al più presto i suoi risultati «primo, perché le idee corrono facilmente dall'uno all'altro che può anticiparne la pubblicazione e, secondo, perché mi sembra vero che molte cose hanno un'epoca nella quale esse sono tro-

vate nello stesso tempo da molte parti, proprio come le violette nascono dappertutto in primavera». Purtroppo passeranno circa dieci anni prima che il lavoro di Janos veda la luce, sicché la paternità della scoperta verrà assunta, nel 1829, dal russo Lobačevskij. Le riflessioni di Janos Bolyai verranno infatti pubblicate nel 1832-33 come appendice di un'opera in due volumi nella quale il padre raccoglie i propri lavori matematici e in particolare i propri tentativi attorno al problema delle parallele: *Tentamen juventutem studiosam in elementa matheseos purae, elementaris ac sublimioris, methodo intuitiva, evidentiaque huic propria, introducendi. Cum appendice triplici* (*Tentativo di introdurre la gioventù studiosa agli elementi della matematica pura, elementare e superiore, secondo un metodo intuitivo, e con l'evidenza ad esso propria. Con tre appendici*).[3]

Il secondo autore cui alludevamo è Friedrich Ludwig Wachter (1792-1817). Visto che, secondo Bolyai, il postulato euclideo dipendeva dal fatto che fosse possibile tracciare una circonferenza per tre punti non collineari, Wachter ritiene che occorra preliminarmente, per poter stabilire una teoria delle parallele, tentare di dimostrare questa esistenza. Egli infatti dà alle stampe nel 1817 una memoria dal titolo *Demonstratio axiomatis geometrici in Euclideis undecimi* (*Dimostrazione dell'undicesimo assioma geometrico di Euclide*),[4] dopo che l'anno precedente ne aveva comunicato il contenuto a

[3] Si osservi che il lavoro di Janos Bolyai non è, come erroneamente di solito si scrive, una delle tre appendici cui si riferisce il titolo. Queste sono inserite regolarmente alla fine del secondo volume, mentre la memoria del nostro autore venne inserita dal padre alla fine del primo volume, con numerazione separata delle pagine (da 1 a 26).

[4] Il diverso numero d'ordine con cui qui ci si richiama al postulato euclideo delle parallele (che viene presentato come undicesimo assioma) deriva dalle varie interpretazioni che si davano alla distinzione fra postulato e assioma. Oltre a quella da noi riportata nella nota 2 a p. 24, Proclo dà almeno altre due interpretazioni diverse di tale distinzione e precisamente: I) un postulato differisce da un assioma come un problema differisce da un teorema (sicché un postulato affermerebbe la possibilità di una costruzione); II) un assioma è diremmo noi oggi «analitico», un postulato è invece una proposizione che pur non essendo un assioma nel senso precedente viene accettato senza dimostrazione (questa distinzione risale ad Aristotele). A seconda quindi del criterio scelto per la distinzione si avrà ovviamente una diversa numerazione delle proposizioni primitive di Euclide. Si noti infine che un'analisi più moderna del postulato euclideo delle parallele ha messo in luce che, sulla base dei precedenti criteri, esso risulta una proposizione di tipo intermedio fra postulati e assiomi, piuttosto che appartenere definitivamente agli uni o agli altri.

Gauss del quale era stato allievo a Gottinga, ove ritiene di aver dimostrato l'assunto a partire dal postulato seguente: quattro punti qualsiasi nello spazio determinano completamente una superficie (la superficie dei quattro punti) e due di queste superfici si intersecano in una linea singola, determinata completamente da tre punti. In effetti in questa sua presunta dimostrazione Wachter non va al di là di elementari considerazioni puramente intuitive, ma fa un'interessantissima osservazione nella succitata lettera a Gauss. Egli afferma infatti che se il quinto postulato fosse falso ci sarebbe una geometria sulla superficie cui tenderebbe una sfera quando il suo raggio tende all'infinito (superficie che nel caso euclideo è appunto un piano) identica a quella dello spazio ordinario. Wachter e Gauss chiamano *antieuclidea* questa geometria e D.M.Y. Sommerville giunge ad affermare che Wachter sarebbe certamente giunto per primo alla scoperta delle geometrie non euclidee se non fosse morto prematuramente.

2.4 *La scoperta delle geometrie non euclidee*

Si è già ripetutamente accennato che nei primi decenni dell'Ottocento, in tutto un clima di rigoroso rinnovamento e di più critica fondazione della matematica, venne a soluzione il problema della indipendenza del quinto postulato euclideo con la costituzione delle geometrie non euclidee. In realtà in questo periodo viene scoperta (ufficialmente con N.I. Lobačevskij e J. Bolyai) *la* geometria non euclidea oggi detta *iperbolica*, ossia il sistema costruito su un'ipotesi corrispondente all'ipotesi dell'angolo acuto di Saccheri o di Lambert: per un punto esterno a una retta si possono condurre (nel piano determinato da quel punto e da quella retta) *due* parallele alla retta data. La scoperta della geometria non euclidea corrispondente all'altra ipotesi saccheriana, quella dell'angolo ottuso, che comporta la non esistenza di rette parallele, avverrà solo nella seconda metà dell'Ottocento e prenderà spunto da una fondamentale memoria del tedesco Bernhard Riemann, da questi letta nel 1854 come dissertazione, ma pubblicata solo nel 1867: *Über die Hypothesen welche der Geometrie zu Grunde liegen* (*Sulle ipotesi che stanno alla base della geometria*). Questo è un caso nel quale – come già avvertito – ci è sembrato opportuno, per dare al lettore una significati-

va visione d'assieme sulle geometrie non euclidee, non attenerci a un rigido criterio cronologico e trattare anche questo secondo sistema, detto oggi di geometria *ellittica* (in questo contesto l'ordinaria geometria euclidea viene anche detta *parabolica*; queste denominazioni risalgono a Felix Klein, 1872). Abbiamo anche incluso nel paragrafo un cenno ai «modelli» delle geometrie non euclidee, essi pure costituiti in realtà nella seconda metà dell'Ottocento; oltre che per il motivo sopra esposto, la loro introduzione a questo punto ci è parsa particolarmente adatta ad illustrare la nuova problematica, di tipo squisitamente logico, che la scoperta delle geometrie non euclidee aveva reso attuale nella ricerca matematica.

2.4.1 *La «Scuola» di Gauss*

Karl Friedrich Gauss[5] è senza dubbio una fra le figure più sconcertanti nella storia della scoperta delle geometrie non euclidee. Comincia infatti a interessarsi del problema della indipendenza del postulato delle parallele mentre è ancora studente a Gottinga, seguendo la via degli studiosi che l'avevano preceduto. Probabilmente anche sotto l'influenza della teoria kantiana dello spazio egli tenta cioè di dimostrare il quinto postulato, dopo aver sottoposto ad acute critiche tutti i tentativi fino ad allora compiuti in questo senso. In un secondo momento egli riconosce l'impossibilità intrinseca di dimostrare il postulato e si convince della legitti-

[5] Karl Friedrich Gauss (1777-1855) fu uno dei più grandi matematici non solo del suo periodo ma di tutti i tempi (venne detto «princeps mathematicorum») e pur occupando una posizione a cavallo fra due secoli così profondamente diversi, come abbiamo visto, per il tenore della produzione matematica, seppe dare notevolissimi contributi nei due sensi, anche se sostanzialmente può considerarsi legato a una mentalità «enciclopedica» settecentesca. Fu professore a Gottinga dal 1807 e si interessò di teoria dei numeri, di algebra, di geometria, di Analisi, di geodesia, di astronomia, di fisica. Fra le sue opere principali: *Disquisitiones arithmeticae* (*Disquisizioni aritmetiche*) del 1801 (nelle quali fra l'altro Gauss introduce la teoria delle congruenze aprendo così la strada allo studio di domini numerici non classici), *Methodus nova integralium valores per approximationem inveniendi* (*Nuovo metodo per trovare per approssimazione i valori degli integrali*) del 1814; *Disquisitiones generales circa superficies curvas* (*Disquisizioni generali sulle superfici curve*) del 1827 (nella quale vengono introdotte e sviluppate le prime nozioni di geometria differenziale, giungendo alla definizione di curvatura ecc.). Di notevolissima importanza scientifica è infine il suo nutritissimo epistolario.

mità della costruzione di un sistema geometrico fondato sulla sua negazione, ossia concepisce chiaramente l'idea di una geometria non euclidea; non pubblica tuttavia il frutto delle sue ricerche per paura delle «strida dei beoti», come ebbe a dire in una lettera a Bessel del 1829, ossia per tema della reazione che l'allora imperante teoria kantiana dello spazio avrebbe determinato contro chi avesse osato mettere in dubbio la natura a priori, necessaria, della geometria euclidea. Quando finalmente si decide a riordinare i risultati delle proprie riflessioni per darli alla stampa, Wolfgang Bolyai gli invia il *Tentamen...* che contiene l'appendice del figlio Janos ove appunto viene presentato, come vedremo, un sistema di geometria non euclidea iperbolica.

Tanto nella prima quanto nella seconda fase del suo atteggiamento a questo riguardo, Gauss ha rapporti diretti o indiretti con gran parte degli studiosi che si interessavano al problema, sicché si può dire che tanto in senso «negativo» (ossia tradizionale) quanto in senso «positivo» (nuove geometrie) egli rappresenta un punto nodale nelle ricerche sulla questione nel periodo cruciale della scoperta dei nuovi sistemi geometrici.

Abbiamo già visto come egli fosse in relazione con Wolfgang Bolyai e con Watcher. Con Gauss era anche in rapporto Bertrand F. Thibaut (1775-1832) autore di un *Grundriss der reinen Mathematik* (*Compendio di matematica pura*) pubblicato in seconda edizione nel 1809, nel quale l'autore ritiene di aver dimostrato il postulato delle parallele nella forma equivalente che la somma degli angoli interni di un triangolo è uguale a due retti. Gauss osserva che nella pseudodimostrazione di Thibaut viene assunta la proposizione da dimostrare, mettendone in evidenza la circolarità.

Direttamente collegati a Gauss (pur lavorando sull'argomento indipendentemente da lui) nel suo periodo «positivo» sono Ferdinand Karl Schweikart (1780-1859) e il nipote di questi Franz Adolf Taurinus (1794-1847) di cui parleremo brevemente più avanti. Va ancora ricordato che fu suo allievo a Gottinga (nel periodo in cui Gauss era ancora convinto della «verità» del quinto postulato) anche Martin Christian Bartels (1769-1836) che poi, nel 1809, sarà insegnante di Lobačevskij all'Università di Kazan.

Naturalmente non è possibile stabilire concretamente quale sia stata la precisa influenza che Gauss eventualmente esercitò su questi studiosi, e va inoltre tenuto presente che in questo periodo

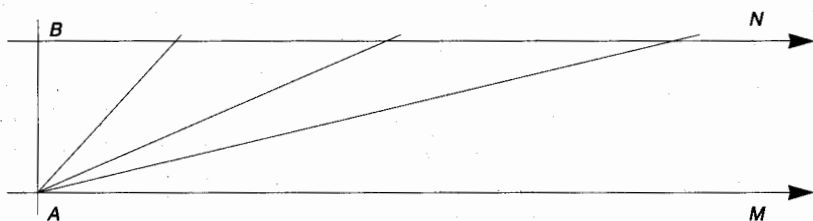
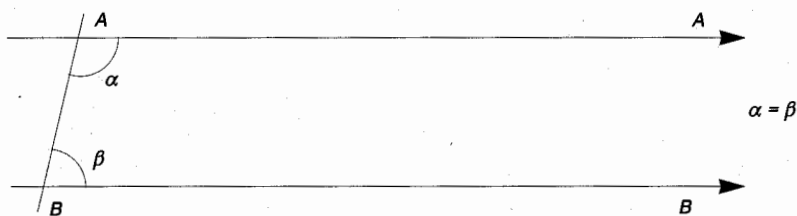
Gauss è considerato in modo indiscusso il più grande matematico vivente, così che era più che naturale che egli costituisse un punto di riferimento obbligato per quanti si cimentavano con l'annosa questione. D'altra parte è anche impossibile valutare con precisione il «debito» di Gauss verso i predecessori, in particolare verso Saccheri e Lambert, dei quali Gauss conosceva e aveva analizzato i tentativi di dimostrazione del quinto postulato. Per poter ricostruire le tappe dell'evoluzione del pensiero di Gauss sulla questione si deve ricorrere al suo epistolario (con Wolfgang Bolyai, Schumacher, Olbers, Gerling, Taurinus e Bessel) nonché a due brevi note pubblicate e a quanto è stato ritrovato nei suoi manoscritti.

Sembra quindi che egli abbia iniziato le proprie «meditazioni» sull'argomento nel 1792, tentando di dimostrare il quinto postulato con lo stesso metodo già usato da Saccheri e da Lambert, ossia assumendolo come falso e cercando di derivare una contraddizione da questa ipotesi. Ancora nel 1804, rispondendo a W. Bolyai che gli aveva sottoposto una propria «dimostrazione» del postulato, pur denunciando l'errore in essa contenuto, Gauss si congratula con l'amico per aver dimostrato in modo definitivo che una possibile direzione di ricerca era in effetti sterile e inconcludente.

Il secondo periodo cui sopra accennavamo comincia per Gauss nel 1813; ora egli sembra aver superato la posizione precedente, abbandonato cioè la speranza di poter dimostrare il postulato, e concepisce già chiaramente l'idea di una geometria completamente diversa da quella euclidea, geometria della quale dimostra numerosi teoremi e che chiama dapprima *antieuclidea*, quindi *astrale* (si veda più avanti Schweikart) e infine geometria *non euclidea*, denominazione questa che ancor oggi si conserva; si convince cioè che la geometria non euclidea, sebbene apparentemente paradossale e del tutto non intuitiva, può riflettere proprietà centrali dello spazio reale. Nel 1831 si decide a raccogliere e forse pubblicare i risultati delle proprie riflessioni, fino allora mantenuti praticamente segreti; ma appunto in quel torno di tempo riceve la comunicazione di W. Bolyai e rinuncia quindi al progetto.

Punto di partenza di Gauss è la seguente definizione di parallelismo:

Se due rette complanari *AM* e *BN* non si incontrano, mentre

d'altra parte ogni retta tracciata per A fra AM e AB incontra BN, allora AM si dice parallela a BN

Gauss dà cioè una definizione di parallelismo diversa da quella euclidea; e si vede subito che se non si accetta il quinto postulato si hanno infinite rette per A che non incontrano BN, ossia sono parallele ad essa in senso euclideo. Gauss dimostra quindi che la relazione così definita è simmetrica, indipendente dalla scelta dei punti A e B sulle rispettive rette, e inoltre che se una retta è parallela ad altre due queste ultime sono parallele fra loro. Successivamente introduce il concetto di *punti corrispondenti* su due parallele AA' e BB' chiamando tali due punti A e B quando la retta AB forma angoli interni uguali con le parallele dalla stessa parte.

Grazie a questo concetto egli può definire la circonferenza come luogo dei punti sulle rette di un fascio[6] corrispondenti a un punto dato. Se le rette del fascio sono parallele (ossia se il sostegno del fascio si allontana all'infinito), nel caso euclideo tale luogo risulta essere una retta (che può quindi essere concepita come una circonferenza di raggio infinito) mentre se non si accetta il quinto postulato (e ci si pone, con Gauss, nella ipotesi dell'angolo acuto) si ottiene come luogo una curva che pur avendo molte proprietà in

[6] Per fascio di rette in un piano si intende l'insieme di tutte le rette del piano che passano per un punto fisso del piano stesso; il punto in questione vien detto «centro» o «sostegno» del fascio. Se il centro di un fascio è all'infinito, il fascio stesso si dice «improprio» ed è costituito da rette parallele fra loro.

comune con una circonferenza euclidea non è tuttavia una circonferenza. Questo luogo avrà una parte importantissima nei sistemi di Lobačevskij e Bolyai e viene oggi detto *oriciclo*.

Gauss riesce a stabilire che nella geometria così costruita esiste una unità di lunghezza assoluta (si ricordino i risultati di Lambert e Legendre) e che nelle sue formule compare una costante k dalla conoscenza della quale dipende la risoluzione di tutti i problemi della geometria non euclidea. In quest'ordine di idee, egli dà per la lunghezza LC della circonferenza di un cerchio di raggio r la formula

$$LC = \pi r \left(e^{\frac{r}{k}} - e^{-\frac{r}{k}}\right)$$

(dove e è la base dei logaritmi naturali e k la costante cui sopra si accennava) e aggiunge che volendo accordare la nuova geometria con l'esperienza quotidiana del nostro spazio fisico, la costante k deve essere posta uguale all'infinito; in altri termini, il sistema geometrico elaborato da Gauss contiene quindi la geometria euclidea come caso particolare per $k = \infty$.

Fra i contemporanei di Gauss direttamente e sicuramente da lui influenzati ricordiamo, come accennato, F.K. Schweikart e il nipote di questi F.A. Taurinus. Schweikart, che era professore di giurisprudenza, pubblica nel 1807 *Die Theorie der Parallellinien nebst dem Vorschlag ihrer Verbannung aus der Geometrie* (*La teoria delle parallele col proposito di eliminarla dalla geometria*) nella quale tratta la teoria delle parallele non in modo indipendente dal quinto postulato, ma fondandola sulla nozione di parallelogramma. Successivamente sviluppa una teoria indipendentemente dalla ipotesi euclidea e nel dicembre 1818 invia a Gauss un «memorandum» di una sola pagina nel quale afferma l'esistenza di due specie di geometrie: la geometria euclidea e una «geometria astrale» nella quale, assunto che la somma degli angoli interni di un triangolo sia diversa da due angoli retti, si dimostra che tale somma è minore di 180°; nota che nelle formule di questa geometria interviene una costante C (che risulta essere il limite superiore dell'altezza di un triangolo isoscele rettangolo) e che la geometria euclidea varrebbe solo nel caso che tale costante C fosse infinita.

Per inciso, la denominazione di «astrale» spiega lo sfondo di queste speculazioni. Come Gauss (e come Lobačevskij più tardi)

Schweikart è convinto infatti che nello spazio reale per distanze astronomiche possa valere una geometria diversa da quella euclidea e che quest'ultima sia la geometria per distanze più piccole. È questo un tipo di speculazione che mostra la natura non formale di queste indagini: l'oggetto è sempre lo spazio reale e l'idea che si fa strada è che la geometria «reale» possa variare in funzione delle distanze secondo una relazione legata in qualche modo al valore delle costanti C e k.

Schweikart non lasciò nulla di pubblicato sull'argomento; Gauss, rispondendo a Gerling nel 1819, loda molto il memorandum di Schweikart e afferma di aver già esteso le idee ivi contenute e che la sua estensione è tale che potrebbe risolvere tutti i problemi solo che fosse noto il valore della costante.[7]

Taurinus fu persuaso dallo zio Schweikart a interessarsi della questione del quinto postulato quando questi, nel 1820, gli sottopose il memorandum già inviato a Gauss e il favorevole giudizio che il grande matematico ne aveva dato. Sembra comunque che egli abbia cominciato a lavorare sull'argomento attorno al 1824, avviandosi tuttavia in una direzione certamente non coincidente con quella sperata dallo zio: essendo convinto della verità del quinto postulato, egli tentò di dimostrarlo. Ovviamente, i suoi tentativi in questo senso fallirono e fu allora che sotto la pressione dello zio e dello stesso Gauss riprese la questione da un altro punto di vista (pur mantenendo inalterata la propria convinzione circa la verità assoluta del quinto postulato euclideo). Nel 1825 pubblica una *Theorie der Parallellinien* (*Teoria delle parallele*) nella quale tratta delle «rette non euclidee» e refuta l'ipotesi dell'angolo ottuso, conducendo inoltre alcune ricerche analoghe a quelle di Saccheri e di Lambert nell'ipotesi dell'angolo acuto. Ritrova in questo caso la costante che già Gauss e Schweikart avevano trovato (egli la chiama «parametro») ma sulla base del fatto che ciò comporterebbe l'esistenza di una unità assoluta di lunghezza refuta tale ipotesi perché contraria alla nostra intuizione dello spazio,

[7] Si noti che la costante C che compare nelle formule di Schweikart non coincide con la k di Gauss (o col «parametro» di Taurinus, si veda più avanti). Le due costanti sono legate dalla relazione

$$k = \frac{C}{\log(1 + \sqrt{2})}$$

pur ritenendola compatibile *logicamente* con i rimanenti postulati euclidei.

L'anno successivo pubblica a Colonia i suoi *Geometriae prima elementa* (*Primi elementi di geometria*) che migliorano le precedenti ricerche e che contengono un'appendice ove egli mostra concretamente come potrebbe essere costruito un sistema di geometria sotto l'ipotesi dell'angolo acuto. A partire da una formula fondamentale della geometria sferica e assumendo il raggio della sfera come immaginario, trova la formula fondamentale di quella che egli chiama la «geometria logaritmico-sferica», nella quale la somma degli angoli interni di un triangolo (in generale minore di due angoli retti) tende a 180° quando i lati del triangolo tendono a zero (si ricordi Lambert). Individua in questa geometria quello che Lobačevskij chiamerà *angolo di parallelismo* e trova la relazione che lega il suo «parametro» (che coincide con la costante k di Gauss) con la costante C di Schweikart.

I risultati di Taurinus confermano, da una parte, certe previsioni di Lambert; Taurinus riconobbe inoltre che la geometria sferica è quel sistema valido nel caso dell'ipotesi dell'angolo ottuso e che l'ordinaria geometria euclidea costituisce una specie di «passaggio» fra la geometria sferica e la propria: il passaggio si determina facendo variare il «parametro» con continuità su tutto il campo reale fino all'immaginario, passando per l'infinito.

2.4.2 *Lobačevskij e Bolyai: la geometria iperbolica*

Abbiamo visto che Gauss, Schweikart e Taurinus erano sostanzialmente giunti a un sistema di geometria non euclidea; un eccessivo scrupolo conformista impedì loro, in particolare a Gauss, di rendere pubblici e imporre i risultati ottenuti, confinando questi ultimi allo stato di geniale esercizio logico e speculativo piuttosto che farli divenire efficace e feconda conquista per la conoscenza umana. È per questo che pur dovendo classificare «di fatto» i tre autori citati fra i fondatori della geometria non euclidea, il merito indiscusso della scoperta va assegnato al russo Nikolaj I. Lobačevskij e all'ungherese Janos Bolyai che non solo giunsero consapevolmente a un sistema non euclideo, ma lo resero altresì pubblico.

Lobačevskij e Bolyai giunsero quasi contemporaneamente, ma in modo sicuramente indipendente, al sistema di geometria iperbolica (che si dimostra coincidere con quelli di Gauss, Schweikart e Taurinus, ossia corrispondere all'ipotesi dell'angolo acuto di Saccheri e di Lambert); nelle pagine seguenti daremo una schematica esposizione di alcuni punti fondamentali della geometria non euclidea iperbolica. A noi infatti non interessa qui tanto dare un'esposizione completa, organica e tecnicamente dettagliata di tale sistema, quanto mettere in luce attraverso alcuni risultati altamente suggestivi e poco intuitivi, da una parte quale sforzo «logico» abbiano compiuto i due autori in esame, dall'altra le motivazioni che li decisero a quel tipo di ricerca e li portarono alla loro scoperta. Di questo secondo aspetto parleremo in particolare nella parte conclusiva di questo paragrafo.

Notiamo intanto subito, con Roberto Bonola, che per quanto riguarda le differenze di impostazione fra i due autori «mentre Lobačevskij dava alla geometria immaginaria [così egli chiamava il suo sistema di geometria non euclidea] un più completo sviluppo, in particolare per quanto riguarda la parte analitica, Bolyai affrontava con maggior completezza la questione della dipendenza o indipendenza dei teoremi della geometria dal postulato di Euclide. Quindi, mentre Lobačevskij cercava principalmente di costruire un sistema di geometria sulla negazione del detto postulato, Janos Bolyai portava alla luce le proposizioni e le costruzioni dell'ordinaria geometria che ne sono indipendenti. Tali proposizioni che egli chiamava *assolutamente vere*, riguardano la *scienza assoluta* dello spazio».

Nikolaj Ivanovič Lobačevskij nacque nel distretto di Makarjev nel 1793. Ancora giovinetto si trasferì a Kazan, dove nel 1809-10 cominciò a frequentare la locale università, fondata appena due anni prima. Nel 1813 divenne *magister*, venne quindi chiamato nel 1816 come professore straordinario e nel 1823 come ordinario nella stessa università, di cui infine, nel 1827, divenne rettore. Venne collocato a riposo nel 1846 e morì nel 1856. Fra il 1809 e il 1813 ebbe a maestro quel Bartels che abbiamo visto essere stato allievo di Gauss a Gottinga. Lobačevskij si interessò della questione del quinto postulato almeno a partire dal 1815, anch'egli tentando in un primo momento di dimostrarlo.

Ebbe tuttavia un'idea precisa della geometria «immaginaria»

attorno al 1823. In quest'anno egli presenta per la pubblicazione sul «Messaggero di Kazan» una memoria sull'argomento che gli viene rifiutata (il manoscritto è stato scoperto solo nel 1898 negli archivi di quella università) e tre anni dopo legge alla sezione matematica dell'università una *Exposition succinte des principes de la géometrie avec une démonstration rigoureuse du théorème des parallèles* (*Esposizione succinta dei principi della geometria con una dimostrazione rigorosa del teorema delle parallele*),[8] purtroppo andata perduta. In questa memoria Lobačevskij illustrava i principi di una geometria più generale della geometria ordinaria, e nella quale per un punto esterno possono essere tracciate due parallele a una retta data, e la somma degli angoli interni di un triangolo è minore di due retti. Queste due prime memorie, con tutta probabilità, non vennero accettate per la stampa perché, come diranno (nel 1949) gli editori delle sue opere, «la teoria di Lobačevskij era incomprensibile per i suoi contemporanei perché sembrava contraddire un assioma basato unicamente su un pregiudizio, consacrato ormai da migliaia di anni».

Nel 1829-30, Lobačevskij pubblica una memoria dal titolo *Sui principi della geometria* che in sostanza riassume, fra l'altro, il contenuto delle conferenze precedenti e successivamente, sempre più desideroso di far conoscere la propria scoperta, pubblica nel 1835 *La geometria immaginaria*, nel 1835-38 i *Nuovi principi di geometria con una teoria completa delle parallele* e nel 1836 *Applicazioni della geometria immaginaria ad alcuni integrali*; quindi, nel 1837 la *Géometrie imaginaire* (*Geometria immaginaria*) e tre anni dopo delle *Geometrische Untersuchungen zur Theorie der Parallellinien* (*Ricerche geometriche sulla teoria delle parallele*). Finalmente nel 1855 dà alle stampe un'esposizione completa della sua teoria sotto il titolo di *Pangéometrie ou précis de géometrie fondé sur une théorie générale et rigoureuse des parallèles* (*Pangeometria o compendio di geometria fondata su una teoria generale e rigorosa delle parallele*). Si può quindi dire con Bonola che la «geometria non euclidea, proprio come fu concepita nel 1816 da Gauss e Schweikart e studiata come sistema astratto da Taurinus nel

[8] Quest'opera venne pubblicata da Lobačevskij in francese. In generale daremo in italiano i titoli delle opere pubblicate originariamente in russo, mentre al solito riferiremo con i titoli originali e relativa traduzione le opere di Lobačevskij pubblicate in altre lingue.

1826, divenne nel 1829-30 una branca riconosciuta del patrimonio scientifico dell'umanità».

Janos Bolyai nacque nel 1802 e fin da ragazzo dimostrò una notevole predisposizione per la matematica. Nel 1818 entrò all'accademia militare di Vienna, donde uscì nel 1823 col grado di sottotenente del genio; nel 1832 è capitano, ma viene collocato a riposo l'anno successivo, perché affetto da una grave forma di ipocondria. Accanto alle ricerche geometriche che dovevano dargli la gloria, Bolyai aveva studiato la possibilità, e vagheggiato l'attuazione, di un linguaggio universale; oltre alla famosa appendice al *Tentamen...* del padre, di cui parleremo più avanti, ha lasciato parti di un volume mai completato (pubblicato nel 1853) dal titolo *Principia doctrinae novae quantitatum imaginariarum perfectae uniceque satisfacientis, aliaeque disquisitiones analyticae et analytico-geometricae cardinales gravissimaeque*; auctore Johan. Bolyai de eadem C.R. austriaco Castrensium pensionato (*Principi di una nuova dottrina delle quantità immaginarie, perfetta e straordinariamente soddisfacente, e altre disquisizioni fondamentali e importantissime, analitiche e analitico-geometriche* di Giovanni Bolyai, pensionato del regio imperial esercito austriaco). Morì nel 1860, quattro anni dopo la morte del padre.

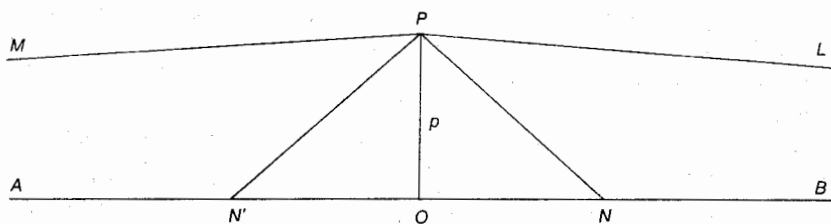
Indirizzato dal padre Wolfgang allo studio della questione relativa al quinto postulato euclideo, egli se ne occupò con fervore nel periodo trascorso all'accademia, ed è certo che almeno fino al 1820 egli accarezzasse l'idea di poter dimostrare il postulato, riuscendo là dove erano falliti altri valenti studiosi. Già nel 1823 aveva tuttavia chiaramente concepito il proprio sistema non euclideo, se scrivendo al padre affermava: «Ho deciso di pubblicare un'opera sulla teoria delle parallele, non appena avrò sistemato il materiale e le circostanze me lo permetteranno. Non ho ancora completato il lavoro ma la strada che ho imboccato mi ha reso quasi certo che, se ciò è in generale possibile, riuscirò nello scopo: questo scopo non è ancora raggiunto ma ho fatto scoperte così meravigliose che ne sono rimasto stupito e che non potrei perdonarmi andassero perdute. Quando le vedrai tu stesso lo riconoscerai. Ora non posso dire di più, aggiungo solo questo: *Ho creato un universo completamente nuovo dal nulla*. Tutto ciò che finora ti ho mandato è null'altro che un castello di carta paragonato a una torre. Sono del tutto persuaso che ciò mi arrecherà onore, come se avessi già completa-

to la scoperta.» Come sappiamo il padre rispose a questa lettera con una calda esortazione, quasi profetica, a pubblicare al più presto il lavoro; Janos comunque mandò al padre un primo estratto del suo lavoro nel 1825, e il manoscritto completo nel 1829. Già sappiamo che il lavoro in questione venne inserito da Wolfgang Bolyai come appendice al primo volume del suo *Tentamen...* del 1832; tale appendice porta il titolo seguente: *Appendix. Scientiam spatii absolute veram exibens: a veritate aut falsitate Axiomatis* XI *Euclidei (a priori haud unquam decidenda) independentem: adjecta ad casum falsitatis, quadratura circuli geometrica.* Auctore ecc. ... Capitaneo. [*Appendice nella quale si presenta la scienza dello spazio assolutamente vera: indipendente dalla verità o falsità dell'undicesimo assioma euclideo (da non decidersi mai a priori): con aggiunta la quadratura geometrica del circolo nel caso della falsità.* Di Giovanni Bolyai, capitano ecc.].

Nello stesso anno il lavoro di Janos venne inviato a Gauss per averne un giudizio. La risposta che sei mesi più tardi Gauss inviava a Wolfgang Bolyai inizia con queste parole: «Se cominciassi dicendo che *non posso lodare quest'opera*, saresti certamente sorpreso per un momento. Eppure non posso fare altrimenti, ché lodarla significherebbe lodare me stesso. Infatti l'intero contenuto dell'opera, la via seguita da tuo figlio, i risultati ai quali egli perviene, coincidono praticamente con le mie meditazioni che, tra le altre, hanno occupato la mia mente negli ultimi trenta o trentacinque anni. Io sono quindi stupefatto». E più avanti, giustificando la sua decisione di non pubblicare nulla sull'argomento, osserva che ciò è avvenuto perché «... la maggior parte delle persone non hanno idee chiare sulle questioni di cui parliamo, e ho trovato ben poche persone che potessero riguardare con un qualche particolare interesse ciò che io gli comunicavo su questo argomento. Per avere un tale interesse è prima di tutto necessario aver studiato con cura la vera natura di ciò che si cerca e su questo quasi tutti sono per lo meno incerti». D'altra parte, dichiarando di aver avuto l'intenzione di sistemare le proprie riflessioni sull'argomento perché «non andassero perdute» con lui, si rallegra che questo ingrato lavoro gli venga risparmiato «proprio dal figlio del mio vecchio amico, che mi ha preceduto in modo così rimarchevole».

Malgrado l'accorta presentazione paterna, l'ipocondriaco Janos reagì violentemente a questa professione di priorità della scoperta da parte di Gauss, e questa lettera determinò in lui una pro-

fonda e definitiva avversione nei riguardi del *princeps mathematicorum.*

Dopo queste brevi note biografiche sugli scopritori della geometria non euclidea iperbolica possiamo passare alla descrizione di alcuni concetti fondamentali di questo sistema. Si considerino una retta *AB* e un punto *P* fuori di essa e si mandi per *P* la perpendicolare *PO* ad *AB*. Si prenda ora un altro punto *N* qualsiasi su *AB*, come mostrato in figura.

Se supponiamo che il punto *N* si allontani sulla retta nel verso, ad esempio, che va da *O* verso *B*, si possono avere due casi:

a) *N* può ritornare al suo punto di partenza dopo aver percorso una distanza *finita*. Ciò corrisponde all'ipotesi saccheriana dell'angolo ottuso e assumendo questa ipotesi come postulato al posto del quinto euclideo (rinunciando ovviamente anche al secondo) si ottiene una geometria non euclidea ellittica, di cui tratteremo nella prossima sezione di questo paragrafo.

b) *N* può continuare a muoversi come sopra detto e la distanza *ON* tendere all'infinito, il che vale nella geometria ordinaria. La retta *PN* tende allora a una posizione limite definita *PL* e *PL* è detta essere parallela a *OB* ($PL \parallel OB$). Ripetendo l'argomento per un punto *N'* preso alla sinistra di *O* si giunge alla conclusione che $PM \parallel OA$.

Ora, nel caso b) si prospettano le due seguenti possibilità:

b1) Nell'ordinaria geometria euclidea le due semirette *PL* e *PM* costituiscono una sola retta, ossia gli angoli $\widehat{OPL}$ e $\widehat{OPM}$ sono entrambi retti (il lettore riconoscerà facilmente che siamo allora nell'ipotesi saccheriana dell'angolo retto).

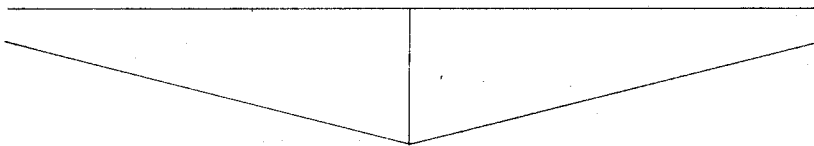
b2) Nella geometria iperbolica invece si assume l'ipotesi che le due semirette *PL* e *PM* siano distinte e che i due angoli sopra detti

siano *acuti* (e il lettore riconoscerà facilmente in questa assunzione l'ipotesi dell'angolo acuto di Saccheri).

Poniamoci ora nel caso b2). Riferendoci alla figura precedente, ne viene che nella geometria iperbolica si avrà la seguente definizione: due rette *PL* e *OB* si dicono *parallele* (in un dato verso o senso; è chiaro infatti da quanto detto sopra che per una retta esistono due possibili direzioni di parallelismo) quando sono complanari, non si incontrano se prolungate indefinitamente, mentre ogni semiretta condotta per *P* entro l'angolo $\widehat{OPL}$ incontra la retta *OB*. Ne viene allora che data una qualunque retta *AB*, per un punto *P* fuori di essa si possono sempre condurre due parallele *PL* e *PM* tali che $PL \parallel OB$ e $PM \parallel OA$. Gli angoli $\widehat{OPL}$ e $\widehat{OPM}$ sono uguali per simmetria; tale angolo, che viene detto *angolo di parallelismo*, dipende solo dalla distanza $PO = p$ e viene indicato con $\Pi(p)$. Le due parallele *separano* per così dire tutte le rette per *P* (ossia il fascio di rette di centro *P*) in due classi: quelle che incontrano la *AB* e quelle che non la incontrano.

Si dimostra che la relazione di parallelismo così definita è indipendente dalla scelta del punto sulla parallela (si ricordi Gauss) e che è simmetrica e transitiva. L'angolo di parallelismo $\Pi(p)$ si dimostra essere una funzione continua di p che diminuisce all'aumentare di p e tende a 0 quando p tende all'infinito, mentre tende a $\pi/2$ al tendere di p a 0. Vale inoltre la relazione

$$\Pi(-p) + \Pi(p) = \pi$$

dove il segno negativo attribuito a p sta a indicare che abbiamo ripetuto il discorso precedente «dall'altra parte» della retta, come indica la seguente figura.

Ne risulta che la distanza fra due rette intersecantisi aumenta indefinitamente, mentre la distanza fra due rette parallele diminuisce e tende a zero nella direzione di parallelismo aumentando

invece senza limite nella direzione opposta: sicché due rette parallele possono considerarsi come intersecantisi all'infinito formando un angolo nullo, ossia uguale a zero. Non è difficile notare che questo era appunto il concetto di retta asintotica intravisto ma non accettato da Saccheri e che inoltre l'analisi compiuta in questo sistema iperbolico consente di separare i due concetti di equidistanza e di parallelismo (chiarendo tra l'altro perché non riuscivano, anzi non potevano riuscire, le «dimostrazioni» del quinto postulato basate sulla definizione di parallelismo come equidistanza). Si dimostra inoltre che proprio ciò che Saccheri aveva ritenuto un fatto «ripugnante alla natura della retta» (perpendicolare comune a due rette all'infinito) è un'interessante conseguenza in questo sistema, nel quale due rette:

1) possono intersecarsi secondo un dato angolo diverso da zero ma non hanno una perpendicolare in comune (ossia sono divergenti o convergenti in senso euclideo); oppure

2) possono non intersecarsi e ammettere una perpendicolare comune (ossia una distanza minima fra loro) senza cioè formare un angolo reale (in altri termini, sono divergenti in entrambe le direzioni); o infine

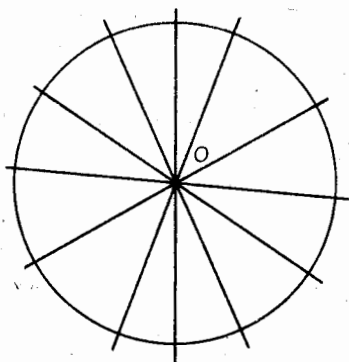
3) possono essere parallele, cioè possono formare all'infinito, dove ammettono una perpendicolare comune, un angolo nullo (ossia sono asintotiche in una direzione e divergenti nella direzione opposta).

È chiaro in particolare che non si hanno rette equidistanti (Lobačevskij definirà il più generale concetto di «curve equidistanti»). Per quanto riguarda l'angolo di parallelismo, si ottiene la formula fondamentale

$$e^{-\frac{p}{k}} = \mathrm{tg}\,\frac{\Pi(p)}{2} \qquad (1)$$

dove e è il numero 2,71..., base dei logaritmi naturali e k una costante caratteristica del sistema.

Altro concetto fondamentale è quello di oriciclo (si ricordi Gauss) che può essere illustrato intuitivamente come segue. Si consideri una circonferenza euclidea di centro O. È noto che essa è perpendicolare a ogni suo raggio ossia è, in linguaggio tecnico che però ha un immediato riscontro intuitivo, la «traiettoria ortogo-

nale» del fascio di rette che ha per sostegno il centro della circonferenza (si veda la figura qui sotto).

Ora supponiamo di far andare all'infinito il centro della circonferenza (ossia, equivalentemente, di aumentare indefinitamente il suo raggio). Le rette del fascio (raggi) diventano parallele e la circonferenza assume una forma limite che nel caso della geometria euclidea è una retta; nel sistema iperbolico tale forma non è una retta bensì una curva uniforme che viene detta appunto *oriciclo* e che resta definita come la traiettoria ortogonale di un fascio di rette parallele. Il concetto si può estendere in modo ovvio dal piano allo spazio (orisfera).

Si possono ovviamente estendere al sistema iperbolico considerazioni trigonometriche, analitiche e proiettive (a queste ultime avremo occasione di accennare brevemente più avanti). Per quanto in particolare riguarda le proprietà metriche, ci limitiamo a ricordare che nella geometria iperbolica la somma degli angoli interni di un triangolo risulta essere minore di due retti, e, detta *difetto* di un triangolo la differenza $D = \pi - A - B - C$ fra un angolo piatto e la somma degli angoli interni del triangolo, si ha che la superficie Δ del triangolo è proporzionale al rispettivo difetto secondo la formula

$$\Delta = k^2 D \qquad (2)$$

dove k^2 è il quadrato della costante già incontrata nella formula (1) e caratteristica della geometria adottata. È chiaro dalla (2) che l'area di un triangolo non può superare il valore πk^2 che è assunto da un triangolo ad angoli tutti nulli, ossia i cui vertici siano tutti all'infinito e i cui lati siano a coppie fra loro paralleli.

2.4.3 *Bernhard Riemann: la geometria ellittica*

Abbiamo già avvertito che avremmo dato per completezza in questo paragrafo un rapido sguardo anche all'altro sistema di geometria non euclidea costituitosi nel XIX secolo, anche se a rigore questo trascende i limiti cronologici che ci siamo posti in questo paragrafo, essendo la scoperta della geometria ellittica un avvenimento della seconda metà dell'Ottocento.

Si è anche accennato che questo nuovo sistema prende le mosse dalle considerazioni contenute nella celebre memoria *Sulle ipotesi che stanno alla base della geometria* che Riemann (1826-66) lesse come dissertazione nel 1854, ma che venne pubblicata solo nel 1867. La memoria di Riemann – una ventina di pagine estremamente dense – va indubbiamente considerata come un punto nodale per la ricerca matematica – e filosofica – sul concetto di spazio. In essa, in particolare, raggiungono un alto grado di rigorizzazione e organicità le considerazioni di geometria differenziale iniziate qualche decennio prima da Gauss (cui Riemann si riferisce esplicitamente) che impostano il problema dei fondamenti della geometria in modo nuovo, originale e profondo.

Riservandoci di ritornare sulla memoria di Riemann nel prosieguo di questo paragrafo ci limitiamo a ricordare la considerazione in essa contenuta che ci permette di presentare succintamente e a livello elementare la geometria ellittica. Si tratta dell'osservazione, sfuggita completamente agli scopritori della geometria iperbolica, relativa alla distinzione fra «illimitatezza» (*Unbegrentzheit*) e «infinità» (*Unendlichkeit*) dello spazio: mentre la prima appartiene all'estensione (si potrebbe dire: è un concetto *qualitativo*) la seconda appartiene alla misura (è un concetto *quantitativo*). Ne risulta che può senza difficoltà concepirsi uno spazio illimitato e pur finito. Orbene, il sistema di geometria ellittica poggia appunto essenzialmente sull'ipotesi che lo spazio sia finito: in particolare ciò si riflette sulla «retta» che, a differenza del caso euclideo e iperbolico, è chiusa e finita.

Sulla base della classificazione fatta nella sezione precedente, ci troviamo evidentemente nel caso a); immediata conseguenza di questa ipotesi è che due rette in un piano si incontrano sempre, anche nel caso in cui siano entrambe perpendicolari a una stessa retta, donde si conclude in particolare che in un piano non posso-

no condursi parallele (in senso euclideo) a una retta per un punto ad essa esterno.

Sotto questa ipotesi si può inoltre dimostrare che tutte le perpendicolari a una data retta *a* e da una stessa banda di essa passano per uno stesso punto *A* che è equidistante da ogni punto della retta *a*. La distanza *AP* dal punto *A* così individuato a un punto qualunque *P* della retta *a* viene detta un *quadrante*. Se ora pensiamo di tracciare tutte le perpendicolari ad *a* dalla banda opposta alla precedente (si ricordi la figura relativa al valore negativo dell'angolo di parallelismo *p* nel caso iperbolico) si viene a individuare un secondo punto *A'* con le stesse caratteristiche di *A*, e si presenta quindi spontaneo il problema di sapere se i punti *A* e *A'* coincidono.

Nel caso infatti che *A* e *A'* siano punti distinti, si dimostra che due rette hanno sempre due punti in comune e si intersecano in una coppia di punti distanti fra loro di due quadranti. Il sistema che così ne risulta (e che come vedremo è intuitivamente assimilabile alla geometria euclidea sulla sfera, se per «rette» si assumono circonferenze massime) viene detto *geometria sferica*.

Nel caso invece che *A* coincida con *A'*, allora due rette si incontrano in un punto e due punti distinti determinano una e una sola retta. È questo secondo sistema che viene tecnicamente chiamato *geometria ellittica*, malgrado di solito si indichi con questa denominazione sia l'uno sia l'altro dei sistemi qui distinti.

A differenza di quanto avviene per il piano come inteso nella geometria euclidea (diremo: piano euclideo) e nel piano iperbolico, il piano ellittico non viene diviso da una sua retta in due regioni distinte: si esprime questa differenza di comportamento dicendo che il piano ellittico ha una «connessione» diversa da quello euclideo o iperbolico. Alla geometria ellittica si possono estendere le usuali considerazioni trigonometriche, analitiche e proiettive; per quanto in particolare riguarda queste ultime, non solo vale anche nel caso ellittico un principio di dualità, ma anzi esso trova in questo sistema la sua più adeguata e generale espressione in quanto si estende oltre che alle proprietà proiettive anche a tutte le proprietà metriche.

Per quanto infine riguarda queste ultime, ci limitiamo a ricordare che nella geometria ellittica la somma degli angoli interni di un triangolo è maggiore di due angoli retti e che la differenza

$E = A + B + C - \pi$ fra questa somma e π viene detta *eccesso* del triangolo; la superficie Δ di un triangolo risulta poi essere proporzionale al suo eccesso secondo la formula

$$\Delta = k^2 E$$

dove k è al solito una costante che caratterizza la particolare geometria ellittica scelta (e che sostanzialmente dipende dalle unità di misura scelte per le lunghezze e gli angoli). Osserviamo infine che anche per la geometria ellittica si possono definire delle «curve equidistanti» dette parallele di Clifford o paratattiche.

2.4.4 *I «modelli» euclidei delle geometrie non euclidee*

I sistemi geometrici non euclidei così costruiti non potevano dirsi incoerenti, nel senso che fra le proposizioni non euclidee concretamente derivate (malgrado esse si presentassero come altamente non intuitive) non si era scoperta alcuna contraddizione logica. È chiaro tuttavia che questa considerazione «sperimentale» oltre a non essere consona allo spirito della ricerca matematica, non garantiva affatto che in un prosieguo di tempo non potessero dimostrarsi proposizioni contraddittorie all'interno dei sistemi in questione. Il problema che a questo punto si poneva in modo naturale, dopo lo sviluppo «tecnico» dei sistemi non euclidei, era quello di dimostrare la correttezza delle nuove geometrie; e questa esigenza, di natura evidentemente metateorica, era indubbiamente inevitabile, una volta che si erano poste tra i principi proposizioni che negavano un postulato non solo «intuitivamente chiaro», ma collaudato da secoli di pensiero matematico.

Affermare la correttezza dei sistemi non euclidei significava allora diverse cose: da una parte mostrarne la plausibilità da un punto di vista *geometrico* individuandone domini d'applicazione, modelli possibili, dall'altra provarne la non assurdità sul piano *logico* il fatto cioè che in questo modo non si giungeva a conclusioni assurde. Ma se l'intuizione è euclidea come è possibile giustificare la geometria non euclidea? È proprio questa difficoltà di fondo che portò in primo piano l'aspetto *formale* dei sistemi geometrici, la loro natura di sistemi ipotetico-deduttivi. Già Lambert aveva osservato come «trattando con la questione si deve ignorare la rappresentazione della materia in oggetto. Poiché i postulati di Euclide e gli assiomi rimanenti sono formulati in parole,

noi possiamo e dobbiamo non fare appello alcuno, nella dimostrazione, al contenuto, ma occorre che la dimostrazione sia sviluppata – se mai è possibile – in un modo puramente simbolico. Da questo punto di vista i postulati di Euclide sono in un certo senso come tante equazioni algebriche da cui dobbiamo ottenere le x, le y, le z, ecc. senza mai voltarci indietro al contenuto in discussione».

D'altra parte lo stesso Lambert sottolineava anche come non potendosi ridurre gli assiomi euclidei a pure formule algebriche, fosse inevitabile considerare rappresentazioni particolari e quindi la ricerca di *modelli* in qualche modo coincideva – sul piano formale se non su quello concettuale – con la ricerca di domini geometrici classici cui applicare – via opportune *reinterpretazioni* dei concetti base – i sistemi non euclidei. Con le parole di Lambert «poiché i postulati non sono formule, possiamo permettere che si traccino figure come guida e ricerca delle dimostrazioni. D'altra parte sarebbe fuori luogo proibire considerazioni e rappresentazioni della materia in oggetto nell'altro aspetto della questione [il problema di trovare assiomi più evidenti di quello euclideo] e richiedere che i nuovi postulati e gli assiomi vengano trovati senza riflettere sul contenuto tirandoli, per così dire, fuori dal cappello». Entrambi gli aspetti della ricerca di correttezza comportavano quindi la costruzione di *modelli*.

Facendo riferimento a dei termini che solo più tardi si sono affermati possiamo infatti dire che il problema della coerenza per un sistema assiomatico può in generale essere affrontato in due modi essenzialmente diversi fra loro:

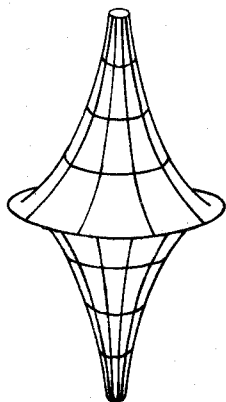
1) si può cercare di dimostrare la coerenza *assoluta* di un sistema, ossia si può tentare di dimostrare che è impossibile in assoluto ottenere una proposizione contraddittoria fra i teoremi del sistema considerato; oppure

2) si può cercare di ottenere una dimostrazione *relativa* di coerenza, riconducendo questa proprietà del sistema considerato alla analoga proprietà di un altro sistema che ci è noto essere coerente (o che comunque si *suppone* essere tale). Con questo tipo di approccio si scarica così sul secondo sistema la «responsabilità» della coerenza del primo. Questo compito viene di solito condotto effettuando una «traduzione», una «rappresentazione» dei concetti del primo sistema in quelli del secondo scelto come riferimento, in modo tale che le proposizioni vere del primo sistema, sottoposte a tale traduzione, risultino proposizioni vere nel secondo sistema. Si dice allora di aver ottenuto un «modello» del primo sistema nel secondo.

Dei due modi sopra prospettati di affrontare il problema della coerenza per sistemi assiomatici, il primo presenta difficoltà intrinseche notevolissime (di cui avremo modo di parlare a lungo nei prossimi capitoli; per ora basterà limitarsi ad accennare che a tutt'oggi non si dispone di una dimostrazione assoluta di non contraddittorietà per nessuna delle teorie interessanti e «sufficientemente potenti», non solo, ma che in una qualsiasi ragionevole accezione di questo «sufficientemente potente» si è dimostrato che è impossibile che ciò avvenga). Il secondo modo invece non urta contro difficoltà di principio e nel caso specifico dei sistemi non euclidei oltre a essere ovviamente applicabile, prospetta quasi come «naturale» scegliere quale sistema di riferimento proprio la geometria euclidea. Se cioè si riesce a ottenere una «traduzione» dei concetti non euclidei in termini di concetti euclidei, e in modo tale che le proposizioni non euclidee vere così tradotte diventino proposizioni euclidee vere, in breve, se si riesce a ottenere un modello euclideo per le geometrie non euclidee, ne rimarrà dimostrata la coerenza di queste ultime, *supposta la coerenza della geometria euclidea* (ipotesi questa che, a livello intuitivo, chiunque si sente di accettare di buon grado). Naturalmente non scenderemo qui in particolari tecnici nella descrizione di questi modelli delle geometrie non euclidee e ci accontenteremo di darne alcune tracce intuitive, limitandoci per semplicità alla planimetria (ossia descrivendo brevemente alcuni modelli euclidei del *piano* ellittico o iperbolico).

In generale si può dire che tali modelli sono stati ottenuti sostanzialmente attraverso due vie, spesso intersecantesi fra loro, e che hanno trovato un'espressione unitaria, a questo riguardo, tramite considerazioni gruppali (per le quali si veda la conclusione di questo paragrafo e il paragrafo successivo): si tratta da una parte della via fondata su considerazioni differenziali che sappiamo sviluppate da Riemann sulle orme di Gauss e che fornì a Eugenio Beltrami (1835-1900) la prima soluzione del problema; dall'altra, della via fondata invece su considerazioni proiettive, sistemate in modo logicamente soddisfacente da Christian von Staudt (1798-1867), portate ad alto grado di sviluppo da Arthur Cayley (1821-95) e sfruttate sotto questo riguardo principalmente da Felix Klein (1849-1925). A Klein in particolare si deve una trattazione unificata della materia tramite la teoria dei gruppi, che venne impiegata anche da Sophus Lie (1842-99) nei contributi del quale in certo senso vengono a convergere considerazioni differenziali, proiettive e gruppali.

Come dicevamo, storicamente la prima interpretazione euclidea

di una geometria non euclidea è quella di Beltrami. Nell'ambito dell'indirizzo differenziale, egli aveva pubblicato nel 1866 una memoria dal titolo *Risoluzione del problema di riportare i punti di una superficie sopra un piano in modo che le linee geodetiche*[9] *vengano rappresentate su rette*; proseguendo questo tipo di ricerche egli pubblicava nel 1868 il celebre *Saggio di interpretazione della geometria non euclidea* nel quale mostrava come potesse ottenersi un modello della geometria iperbolica «rappresentando» (almeno per quanto riguarda porzioni limitate del piano non euclideo) gli elementi fondamentali non euclidei «punto», «retta», ecc. su opportuni elementi di una opportuna superficie euclidea, detta *pseudosfera*, caratterizzata dall'avere curvatura[10] costante e negativa (si veda la figura qui sotto). In particolare, in questa «traduzione» ai «punti» non euclidei corrispondevano punti

[9] Per linea «geodetica» o semplicemente «geodetica» di una superficie si intende una linea della superficie che goda dell'analoga proprietà delle rette nel piano euclideo di rappresentare il cammino minimo (la minima distanza) fra due punti del piano. Ovviamente le geodetiche del piano euclideo sono le rette, le geodetiche di una superficie sferica le circonferenze massime. Per una qualunque superficie si può dare un'espressione analitica generale per le relative geodetiche.
[10] Nel contesto nel quale ci muoviamo ci limiteremo a dare una chiarificazione intuitiva di questo termine, che peraltro denota un concetto fondamentale di tutta l'impostazione differenziale di Gauss e Riemann. In questo senso diremo che la curvatura di una linea in un suo punto è un *numero* che «misura» di quanto la linea, in quel punto, si scosta dall'andamento rettilineo (o altrimenti detto dalla sua tangente in quel punto); analogamente, la curvatura di una superficie in un suo punto è il *numero* che misura di quanto la superficie, in quel punto, si scosta dall'andamento piano (dal suo piano tangente in quel punto). Si comprende che in generale la curvatura di una superficie varierà da punto a punto. Casi particolarmente interessanti costituiscono quelle superfici la cui curvatura si mantiene costante in ogni loro punto; tra di essi esempi intuitivamente comprensibili sono

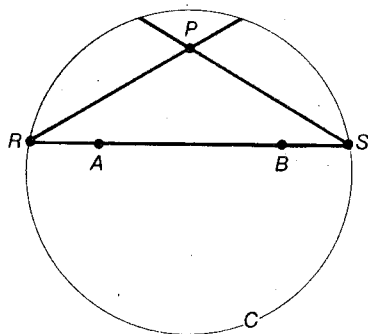
sulla pseudosfera, alle «rette» non euclidee linee geodetiche sulla pseudosfera e infine al «piano» (o più precisamente a una porzione del piano) non euclideo la (una porzione della) superficie della pseudosfera. Tralasciamo qui di precisare le corrispondenze fra i concetti di «distanza» e di «angolo» non euclidei e gli analoghi concetti euclidei; vedremo d'altronde che essi rientrano in una classificazione generale che delle varie geometrie può darsi in termini proiettivi.

Come mostrato da Klein, si può dare un modello euclideo molto più intuitivo per la geometria iperbolica (tralasciando al solito distanze e angoli) fissando sul piano euclideo una conica C (che noi per comodità assumeremo essere una circonferenza) e stabilendo il seguente «dizionario» per la traduzione dei più elementari termini iperbolici nei corrispondenti termini euclidei.

Geometria iperbolica	Geometria euclidea
punto	punto interno a C (esclusi i punti del contorno)
retta	ogni segmento aperto RS di secante, ossia ogni corda di C estremi esclusi
piano	l'insieme dei punti e delle rette sopra definiti
⋮	⋮

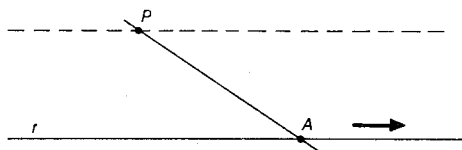
Si verifica subito che per due punti passa una retta e una sola (per due punti interni a C passa una corda e una sola) che due rette hanno in comune uno e un solo punto (due corde di C si incontrano in uno e un solo punto di C) ecc. Si può cioè verificare che in questo modello valgono gli assiomi[11] di una qualsiasi sistemazione

il piano, la cui curvatura è costante e uguale a zero in ogni punto; e la superficie sferica per la quale è evidente la costanza della curvatura. Se si conviene di considerare «positiva» la curvatura costante della sfera, si chiamerà «negativa» la curvatura costante di superfici, come appunto quella della pseudosfera, che è per così dire incurvata «in senso opposto» a quella della sfera.

[11] Ci si convince facilmente che per stabilire che una data «traduzione» è modello di una teoria posta in forma assiomatica, è sufficiente verificare che sotto tale

assiomatica della geometria euclidea (e ciò è vero anche per quelli che contengono i corrispettivi dei concetti di distanza e angolo che non abbiamo sopra considerato) con la sola eccezione del postulato delle parallele. Si vede infatti che dati una retta del nostro modello (ad esempio la corda *RS*) e un punto *P* fuori di essa, esistono *due* parallele (e anzi infinite in senso euclideo) alla retta data (sono le rette *PS* e *PR* che incontrano la *RS* nei punti *R* e *S* che tuttavia *non* sono punti «propri» del nostro modello; ad esse vanno aggiunte tutte le rette interne ai coni segnati in grigio nella figura precedente).

I concetti metrici fondamentali di distanza e di angolo si ottengono come dicevamo con considerazioni proiettive, una volta che si sia ampliato il piano con l'aggiunta di punti «impropri» o «all'infinito»[12] (piano proiettivo). L'insieme dei punti impropri del piano costituisce quello che Cayley chiama l'*assoluto* del piano. Servendosi di quest'ultimo concetto si può allora dimostrare che l'intera geometria metrica è determinata dalla natura dell'assoluto, nel senso che essa è null'altro che la geometria proiettiva in re-

interpretazione diventano veri gli *assiomi* della teoria in questione (nel senso che da questo deriva immediatamente la verità dei teoremi).

[12] Si consideri la semplice situazione illustrata in figura. Se il punto *A* continua a

spostarsi nel verso indicato dalla freccia, appare chiaro che quando la retta *PA* sarà divenuta parallela alla *r* (posizione tratteggiata in figura) il punto *A* si sposta all'infinito sulla retta *r*. Ciò si può esprimere dicendo che la retta *r* e la sua parallela così individuata (e quindi tutte le sue parallele) hanno in comune un pun-

lazione all'assoluto, se si definiscono la distanza di due punti A e B e l'angolo fra le due rette p e q rispettivamente mediante le espressioni

$$\text{dist } AB = K\log(ABRS), \qquad \text{ang } pq = k\log(pqrs) \qquad (3)$$

ove le notazioni $(ABRS)$ e $(pqrs)$ indicano il birapporto[13] dei quattro punti A, B, R, S in quest'ordine o delle quattro rette p, q, r, s in quest'ordine. Riferendo le formule (3) alla figura precedente (e ne vedremo subito la giustificazione) si vede che per determinare la distanza fra due punti A e B si fa il birapporto di questi e dei due punti R, S nei quali la retta per AB incontra la C; per determinare l'angolo fra due rette P e q si fa il birapporto delle due rette e delle tangenti r e s alla C dal loro punto d'incontro.

Per quanto riguarda la natura dell'assoluto, si dimostra intanto che esso è una conica e si ottengono le seguenti caratterizzazioni: se esso è una conica reale non degenere si ha la geometria iperbolica (sicché nel nostro modello precedente, e nella nostra figura, C era l'assoluto del piano iperbolico); se esso è una conica degenere si ha l'ordinaria geometria euclidea (parabolica); se infine l'assoluto è una conica immaginaria non degenere si ha la geometria

to all'infinito, il che è equivalente a dire che hanno una comune *direzione*. Orbene, per punto *improprio* o all'infinito si intende appunto la direzione individuata da una retta e da tutte le sue parallele. Se si conviene di ampliare il piano considerando accanto ai punti «propri» l'insieme di tali punti «impropri» si potrà dire ad esempio che due rette si incontrano sempre, in un punto proprio (che corrisponde alla normale accezione «dell'incontrarsi» di due rette) o in un punto improprio (che quindi traduce in termine di «intersezione» la comune accezione di parallelismo). Si noti che su ogni retta si hanno due punti all'infinito che nel caso euclideo coincidono, nel caso iperbolico sono invece distinti.

[13] Si dice birapporto di quattro punti allineati A, B, R, S, e si indica con $(ABRS)$ il *numero* dato del doppio rapporto

$$(ABRS) = \frac{(ABR)}{(ABS)} = \frac{AR}{BR} : \frac{AS}{BS}$$

dove con AR, BR, ecc., si intendono le distanze, misurate in questo ordine, dei punti A e R, B e R, ecc. a partire da una comune origine.
Analogamente, per quattro rette p, q, r, s di un fascio si ha

$$(pqrs) = \frac{\text{sen }\widehat{pr}}{\text{sen }\widehat{qr}} : \frac{\text{sen }\widehat{ps}}{\text{sen }\widehat{qs}} .$$

Il birapporto è uno degli invarianti proiettivi fondamentali, resta cioè invariato per proiezioni e sezioni.

ellittica. In ogni caso i punti dell'assoluto sono a distanza infinita da tutti gli altri punti del piano e le tangenti all'assoluto formano un angolo infinito con tutte le altre rette.

Se ora, per così dire, si rovescia la situazione e si parte *definendo* distanze e angoli attraverso le formule (3), si ottiene un sistema geometrico generale che comprende come casi particolari le tre geometrie ellittica, iperbolica e parabolica; e la natura della particolare geometria viene ad essere determinata dai valori delle costanti K e k che in esse figurano. In generale il valore di queste costanti dipende soltanto dalle rispettive unità di misura scelte per distanze e angoli; ma si ha una distinzione essenziale a seconda del «tipo» dei valori scelti, e precisamente: per ogni valore *reale* di K (considerazioni analoghe valgono per k) si ottiene una geometria iperbolica; per ogni valore *immaginario* di K si ottiene un sistema di geometria ellittico; se infine si pone K uguale all'*infinito* si ottiene la ordinaria geometria euclidea parabolica.

Sulla base delle considerazioni precedenti la costruzione di un modello per le geometrie non euclidee diventa, nel caso particolare della geometria ellittica, una questione analitica abbastanza complessa. Per questo tipo di geometria si può però dare come per il caso iperbolico un modello sufficientemente perspicuo e intuitivo considerando il seguente «dizionario»:

Geometria ellittica	Geometria euclidea
punto	coppia di punti diametralmente opposti su una prefissata superficie sferica [se si assumesse semplicemente: punto di una opportuna superficie sferica si otterrebbe un modello della geometria sferica]
retta	circonferenza massima della data superficie sferica
piano	la superficie sferica considerata
.	.
.	.
.	.

Anche in questo caso si verifica immediatamente che valgono gli assiomi euclidei, eccetto, naturalmente, il secondo (sull'infinità della retta) e il quinto (non si hanno rette parallele).

2.4.5 *Sguardo conclusivo sulle geometrie non euclidee*

Secondo Felix Klein nello sviluppo delle geometrie non euclidee è possibile distinguere tre periodi. Il primo, che comprende l'opera di Gauss, Lobačevskij e Bolyai, è caratterizzato dall'impiego del metodo sintetico, che applica metodi di geometria elementare; il secondo, che è collegato alla rappresentazione geodetica e applica metodi della geometria differenziale, comprende i contributi di Riemann, Helmholtz, Lie e Beltrami; il terzo, infine, che tratta della rappresentazione proiettiva, applica cioè i principi della geometria proiettiva, è iniziato da Cayley e comprende lo stesso Klein. A questi periodi Sommerville propone di aggiungerne un quarto, relativo alla ricerca sui fondamenti logici della geometria fondata su un dato insieme di assiomi. A parere di Sommerville questo ultimo periodo «è stato inaugurato da Pasch, ma per ritrovarne i veri inizi dobbiamo risalire a von Staudt. Tale periodo comprende Hilbert e la scuola italiana rappresentata da Peano e Pieri; in America il principale rappresentante è Veblen».

La distinzione precedente è particolarmente utile per quanto riguarda la nostra esposizione, soprattutto per il primo e il quarto periodo in essa individuati. Il primo periodo infatti mette in luce implicitamente un aspetto della «rivoluzione» provocata dai fondatori delle geometrie non euclidee, che non riguarda tanto l'affi-

namento e il potenziamento di strumenti tecnici, quanto un vero e proprio mutamento di punto di vista, di atteggiamento generale, di idee: i risultati da essi ottenuti sono infatti fondati su uno strumento matematico sostanzialmente coincidente con quello euclideo, in altri termini, paradossalmente, anche Euclide avrebbe potuto *tecnicamente* dare origine alle... geometrie non euclidee. Che l'idea di alternative al quinto postulato si fosse presentata già prima di Euclide e ad Euclide stesso, e non sia peregrina, è stato provato dalle ricerche di I. Toth cui abbiamo già fatto riferimento. Il quarto periodo mostra invece come tutta la questione relativa a queste geometrie sfoci in modo naturale in una tematica di natura squisitamente ed essenzialmente *logica*, inserendosi di diritto fra gli avvenimenti scientifici che hanno maggiormente stimolato questa tematica.

Se ora consideriamo non solo e non tanto i contributi «tecnici» dei diversi studiosi a cui spetta il merito di aver costituito questi nuovi sistemi geometrici, ma allarghiamo il discorso sino a comprendere anche il loro atteggiamento generale nei riguardi del problema in questione, non è difficile scorgere in tali atteggiamenti delle singolari coincidenze in relazione al problema della geometria come descrizione dello spazio. Queste coincidenze si presentano indubbiamente con sfumature assai diverse e con diversi gradi di consapevolezza, ma sono sostanzialmente concordi sulla considerazione a posteriori della geometria. In altri termini questi vari autori tendono tutti a smentire l'affermazione kantiana secondo la quale gli assiomi della geometria sono conseguenze necessarie di una forma trascendentale, data a priori, della nostra facoltà intuitiva.

In questo senso infatti, a nostro parere, vanno intese già le parole di J. Bolyai quando questi, risolvendosi a non più tentare di *dimostrare* il postulato delle parallele, scrivendo al padre si dichiarava convinto «che non bisogna forzare la natura, né conformarla a nessuna chimera ciecamente costituitasi; che, d'altra parte, si deve riguardare la natura ragionevolmente e naturalmente, come si farebbe con la verità, e accontentarsi soltanto di una rappresentazione di essa che ne differisca il meno possibile». È con questa convinzione che J. Bolyai impiega sì il metodo deduttivo, ma senza decidere a priori della verità o falsità del quinto postulato, ossia elabora di fatto un *sistema ipotetico-deduttivo.*

Ancor più deciso e evidente l'atteggiamento di Lobačevskij quando nei *Nuovi principi di geometria* del 1835 afferma: «I vani sforzi compiuti dai tempi di Euclide, per il corso di duemila anni, mi spinsero a sospettare che nei concetti stessi della geometria non si racchiuda ancora quella verità che si voleva dimostrare, e che può essere controllata, in modo simile alle leggi fisiche, soltanto da esperienze quali, ad esempio, le osservazioni astronomiche. Essendo infine convinto della verità della mia congettura, e giudicando completamente risolto il difficile problema, scrissi su di esso una comunicazione nel 1826.» È qui abbastanza chiaramente avvertibile una concezione dello spazio che in opposizione alla teoria kantiana (spazio come intuizione trascendentale, presupposto necessario di ogni esperienza) ritrova il proprio fondamento su considerazioni empiriche, portando la geometria fra le scienze sperimentali.

Ancora più esplicita e consapevole è la critica di Riemann. Riconosciuto, nella memoria più volte citata, come gli elementi e i concetti fondamentali della geometria fossero ancora ai suoi tempi avvolti nell'oscurità, egli si propone come compito di esplicitare un concetto generale di «grandezza molteplicemente estesa» (varietà pluridimensionale) a partire dal concetto generale di grandezza. Da ciò risulterà che «una varietà n-dimensionale è suscettibile di diverse determinazioni di misura [*metriche*, noi diremmo oggi] e che lo spazio [fisico, della nostra esperienza] non è che un caso particolare di varietà tridimensionale». Da questa impostazione scende come conseguenza necessaria che «le proposizioni della geometria non si possono derivare da concetti generali di grandezza ma che quelle proprietà in base alle quali lo spazio si differenzia da altre pensabili varietà tridimensionali, possono essere assunte sulla base dell'esperienza. Da qui sorge il compito di elencare i fatti più semplici sulla base dei quali si possono determinare le metriche dello spazio, un compito che per sua stessa natura non è completamente determinato; infatti si possono dare svariati sistemi di cose semplici che sono sufficienti alla determinazione delle metriche dello spazio... Questi fatti, come del resto tutti i fatti, non sono necessari, ma posseggono soltanto una certezza empirica, sono delle ipotesi; si può quindi indagare la loro probabilità, che nell'ambito dei limiti dell'osservazione è molto grande, e da qui giudicare sulla liceità di una loro estensione al di là dei limiti

dell'osservazione, tanto nell'infinitamente grande quanto nell'infinitamente piccolo».

Precisato quindi il concetto di varietà *n*-dimensionale come aggregato i cui elementi (punti) si possono determinare con l'assegnazione di *n* grandezze variabili (coordinate), Riemann passa appunto a considerare le possibili metriche applicabili a una tale varietà sotto l'unica ipotesi che le linee della varietà stessa posseggano una lunghezza indipendente dalla posizione, sicché ogni linea possa venir misurata da ogni altra. Sotto queste ipotesi Riemann deve allora esplicitare la dipendenza della lunghezza di una linea che «parta» da un punto, dai corrispondenti differenziali delle coordinate del punto; in altri termini, un punto P della varietà può essere determinato da n coordinate $x_1, \ldots, x_n$, e se $x_1 + dx_1, \ldots, x_n + dx_n$ sono le coordinate di un altro punto Q della varietà a distanza infinitesima da P, occorre dare la lunghezza ds dell'elemento lineare PQ in termini degli incrementi $dx_1, \ldots, dx_n$. A questo punto Riemann fa l'ipotesi che il ds sia uguale alla radice quadrata di una funzione omogenea di secondo grado dei differenziali delle coordinate. Egli giustifica quest'ipotesi restrittiva argomentando in modo euristico, presentandola come la più semplice fra tutte quelle che possono soddisfare il problema e dichiarando peraltro esplicitamente che il ds può essere dato mediante espressioni più complesse di quella da lui adottata. In vista quindi della successiva applicazione delle sue argomentazioni allo spazio reale egli fornisce esplicitamente la formula per la metrica, ossia per il ds, nel caso di varietà n-dimensionali a curvatura costante α.

Nel corso dell'applicazione allo spazio fisico Riemann introduce la distinzione cui sopra abbiamo accennato fra illimitatezza e infinità dello spazio, e dopo averla enunciata afferma che il fatto «che lo spazio sia una varietà tridimensionale illimitata è un'ipotesi che viene assunta da ogni concezione del mondo esterno» che sulla base di continue applicazioni «viene continuamente confermata. L'illimitatezza dello spazio possiede così una maggiore certezza empirica di ogni altra esperienza esterna. Ma da essa non segue assolutamente l'infinità: anzi, se si assume l'indipendenza dei corpi dalla posizione e allo spazio si assegna inoltre una curvatura costante, esso dovrebbe necessariamente essere finito non appena la misura di questa curvatura avesse un valore positivo sufficientemente piccolo». La memoria riemanniana si chiude con l'impor-

tantissima osservazione che «una ricerca la quale voglia condurre a un'effettiva conoscenza dello spazio e delle sue proprietà dovrebbe svolgersi nell'ambito [non della geometria ma] di un'altra scienza, nell'ambito della fisica...».

La problematica relativa all'indagine sulla natura dello spazio fu argomento attuale e tipico della seconda metà dell'Ottocento (e prepara direttamente, da più punti di vista, la sistemazione assiomatica hilbertiana della geometria che si avrà attorno al 1900). I brevissimi cenni con i quali concluderemo questo paragrafo hanno il solo scopo di mostrare insospettati collegamenti fra questo tipo di ricerche, in gran parte propiziate come si è visto dalla scoperta delle geometrie non euclidee, e le ricerche algebriche che nella seconda metà dell'Ottocento vanno moltiplicandosi con rinnovata vitalità sulla scorta di nuove concezioni che anche in questo campo erano state introdotte nella prima metà del secolo e delle quali ci occuperemo nel prossimo paragrafo.

Nel 1868 Hermann von Helmholtz (1821-94) pubblica la memoria *Über die Tatsachen, die der Geometrie zu Grunde liegen* (*Sui fatti che stanno alla base della geometria*) il cui stesso titolo mostra già una singolare affinità con le ricerche contenute nella memoria di Riemann. Il lavoro di Helmholtz, pur essendo del tutto indipendente, viene in effetti a porsi in diretta connessione con quello di Riemann, assumendo in particolare l'ipotesi fondamentale che uno spazio a *n* dimensioni sia una varietà *n*-dimensionale nel senso di Riemann. Accanto a questa ipotesi Helmholtz assume inoltre l'esistenza di corpi rigidi in natura,[14] la libera mobilità degli stessi[15] e la monodromia dello spazio.[16] Su questa base la relazione fra la memoria di Riemann e quella di Helmholtz ci viene chiarita da quest'ultimo come segue: «Le mie ricerche si distinguono dalle ricerche di Riemann per il fatto che io ho esaminato più da presso l'importanza della condizione restrittiva su detta[17] – onde lo spa-

[14] «... com'è necessario per poter intraprendere il confronto delle grandezze spaziali mediante una verifica di congruenza».

[15] «... ovvero si presuppone che ogni loro punto possa trasferirsi in modo continuo al posto di qualsiasi altro punto, fino a che il punto stesso non sia vincolato dalle equazioni sussistenti fra esso e tutti gli altri punti del sistema rigido».

[16] «... che è analoga alla monodromia delle funzioni di una grandezza complessa e che si esprime nel fatto che due grandezze congruenti rimangono tali seppure una di esse abbia subito un rivolgimento secondo un qualsiasi asse di rotazione».

[17] Relativa alla particolare espressione per il *ds* scelta da Riemann.

zio reale si differenzia da altre varietà pluridimensionali – quando si voglia motivare quello che è il fulcro di tutta l'indagine, ossia il principio secondo il quale il quadrato dell'elemento lineare è una funzione omogenea di secondo grado dei differenziali delle coordinate. Si può dimostrare che, ammettendo sin dall'inizio un'illimitata libera mobilità di figure per sé rigide in tutte le parti dello spazio senza modificazioni di forma, l'ipotesi originaria di Riemann può essere ricavata come conseguenza da premesse molto meno ristrette. Io presi le mosse dal concetto che ogni misura primitiva dello spazio dipende da un'osservazione di congruenza.»

In effetti le argomentazioni di Helmholtz, oltre a non essere debitamente rigorose da un punto di vista matematico, sostituivano all'ipotesi di Riemann da una parte l'ipotesi non più plausibile e comunque non certo verificata dell'esistenza *in natura* di corpi rigidi (i quali viceversa non sono che frutto di una pura astrazione matematica) e d'altra parte, con la «libera mobilità» di questi corpi, veniva in certo senso a postulare una sorta di spazio assoluto.

L'intera questione venne ripresa dal matematico norvegese Sophus Lie, al cui pensiero si era ispirato Klein nel suo Programma, in due fasi successive nel 1890 e nel 1893. Lie, fondandosi su considerazioni gruppali che estende al caso continuo, ritiene di poter evitare le ipotesi helmholtziane riducendo il concetto di movimento a quello di trasformazione fra sistemi di coordinate e quello di congruenza all'invarianza rispetto a tali trasformazioni. L'idea di Lie è quella di considerare gruppi continui di trasformazioni infinitesime, e di porre alla base delle sue considerazioni un corrispondente assioma di «libera mobilità nell'infinitesimo». Con questa impostazione, a partire dal gruppo delle trasformazioni proiettive, Lie ritrova il concetto di assoluto e ne conclude che i movimenti dello spazio costituiscono un sottogruppo del gruppo generale delle trasformazioni fra punti che lascia invariato l'assoluto, stabilendo come conseguenza di questo risultato che gli unici tipi possibili di geometria metrica sono appunto l'iperbolica, la parabolica e l'ellittica.

Oltre quindi a notare come si venga stabilendo un armonico e stretto collegamento fra rami della matematica apparentemente così distanti fra loro, come considerazione conclusiva possiamo osservare che il problema che in realtà veniva posto qui in discus-

sione non era tanto e soltanto quello dei fondamenti della geometria, quanto quello, di dimensioni ben più generali, della natura stessa dello spazio. Tale problema, tradizionalmente di pertinenza della speculazione filosofica, viene ora affrontato da un punto di vista scientifico ed è chiaro che esso non perde così, certamente, la pregnanza che aveva avuto nel passato, ma anzi acquista una dimensione più completa e feconda proprio a causa delle nuove e sottili determinazioni che ad esso vengono assegnate da questo diverso e potente tipo di approccio. Se questo poneva in primo piano il problema di come determinare la natura dello spazio *reale*, imponeva d'altro canto considerazioni nuove sullo status della geometria e della matematica in generale come *teoria*. Su che cosa si fonda la validità della inferenza matematica? Come si possono studiare i rapporti tra assiomi e proposizioni geometriche senza presupporre legami che hanno giustificazione solo intuitiva in riferimento a interpretazioni che non sono le uniche possibili? In altre parole, come è possibile analizzare sistematicamente i rapporti tra inferenza e possibili interpretazioni?

3. L'EVOLUZIONE DELL'ALGEBRA NEL CONTINENTE E IN INGHILTERRA

Gli interrogativi precedenti si situano in un contesto che vede in ogni campo della matematica l'affermarsi di una tendenza alla generalizzazione, alla conquista di nuovi domini d'applicazione, con una conseguente necessità di revisione critica, di rigore, sconosciuta nel Settecento. Come già avvertito, ci limiteremo qui ad analizzare l'evoluzione dell'algebra tanto nel Continente quanto in Inghilterra, rilevandone, ove possibile, analogie e differenze e tentando di mettere in luce l'influenza che il discorso algebrico in modo speciale ebbe nella preparazione dell'ambiente inglese che dovrà assistere a una sì rigogliosa rinascita della logica formale. Come considerazione generale di fondo si può intanto osservare che ciò che rende possibile e concretamente concorre a instaurare lo sviluppo moderno della logica formale è proprio il fatto che in questo periodo comincia ad avvertirsi con chiarezza il carattere non necessario dell'ancoramento della disciplina matematica a modelli privilegiati e precostituiti, quali quello geometrico o arit-

metico (numerico) e si assiste, di conseguenza, a una progressiva problematizzazione del rapporto tra sviluppo deduttivo e applicazione, momento formale e interpretazione.

Questo fenomeno, pur in gradi e con sfumature diverse per le varie discipline matematiche e per i vari autori (o addirittura per i vari paesi) si prospetta come del tutto generale. Nelle pagine dedicate all'evoluzione dell'algebra in Inghilterra vedremo che in effetti il motivo dominante donde prende lo spunto concreto la nuova impostazione della logica formale è costituito dalla dimensione del tutto caratteristica assunta nell'isola dalla ricerca algebrica.

3.1 *L'algebra nel Continente*

Dopo la fioritura del XVI secolo si era avuto in campo algebrico un ristagno nella ricerca soprattutto perché il problema fondamentale era divenuto – dopo che gli algebristi italiani del Cinquecento avevano dato le formule risolutive per radicali delle equazioni algebriche generali dei primi quattro gradi – quello di trovare un'analoga formula per le equazioni di grado superiore al quarto. Già verso la fine del Settecento tuttavia si presagiscono i primi sintomi della necessità di impostare questo problema in forma diversa o, altrimenti detto, si comincia a intuire che la soluzione del problema non si otteneva per il semplice motivo che il problema stesso era mal posto. Fra il 1770 e il 1771 Lagrange aveva pubblicato la memoria *Réflexions sur la résolution algébrique des équations* (*Riflessioni sulla risoluzione algebrica delle equazioni*) nella quale eseguiva un confronto dei vari metodi impiegati nella risoluzione delle equazioni di grado uguale o minore al quarto e tentava di spiegare perché tali metodi non possono essere estesi già alle equazioni di quinto grado; egli notava infatti che nei primi quattro casi il problema si può riportare a scrivere un'equazione *risolvente* che risulta di grado inferiore alla data, mentre se il grado n dell'equazione è maggiore di 4 ciò non avviene, anzi la risolvente in questione risulta essere di grado maggiore all'equazione di partenza. La questione lo portava a uno spunto di considerazioni gruppali, nel senso che Lagrange viene condotto a considerare delle funzioni razionali delle radici delle equazioni e il loro comportamento rispetto alle permutazioni delle radici stesse.

Orbene, anche questo annoso problema viene risolto nella prima metà dell'Ottocento, essenzialmente ad opera di due matematici geniali e giovanissimi, accomunati, oltre che dalla gloria scientifica, dall'aver avuto entrambi una vita travagliata e tormentata (anche se per motivi assai diversi) e troncata in età ancor giovanissima. Alludiamo al francese Evariste Galois e al norvegese Niels Henrik Abel, di cui ci occuperemo brevemente in questo paragrafo.

Per esattezza storica va tuttavia ricordato che il primo autore che pervenne alla dimostrazione dell'impossibilità di risolvere per radicali (nel caso generale) le equazioni algebriche di grado superiore al quarto fu il medico italiano Paolo Ruffini (1765-1822) che pur esercitando la sua professione si occupa intensamente di matematica che insegna anche, per un certo periodo, all'università di Modena. Il risultato in questione venne da lui pubblicato nel 1798 nella memoria *Teoria generale delle equazioni* che egli inviò nel 1810 all'Institut di Parigi, senza tuttavia ottenere alcuna risposta dalla commissione nominata per esaminarla (Lagrange, Legendre e Lacroix); analogo esito ebbe l'invio della memoria alla Royal Society. In particolare, Poisson trovò «troppo vaga» la dimostrazione di Ruffini, che oggi tuttavia ha ricevuto generale riconoscimento e che contiene sicuri elementi di applicazione della teoria delle sostituzioni, dalla quale si svilupperà la teoria dei gruppi. La dimostrazione di Ruffini presentava in effetti lacune e oscurità; per il suo contributo egli si era ispirato alle considerazioni di Lagrange sulla teoria delle permutazioni delle radici delle equazioni e, diremmo oggi, era giunto a far dipendere la risolubilità dell'equazione stessa dall'esistenza di particolari sottogruppi del gruppo di tali permutazioni.[18]

[18] Sarà opportuno a questo punto ricordare una definizione di «gruppo» come inteso nell'algebra astratta moderna. Dicesi dunque «gruppo» un insieme G non vuoto di elementi qualsiasi fra i quali sia definita una legge di composizione interna (un'operazione) che indicheremo per semplice giustapposizione, e tale che
a) l'operazione sia associativa, vale a dire per ogni elemento x, y, z di G si abbia $(xy)z = x(yz)$;
b) G possiede un elemento neutro e, vale a dire un elemento per il quale risulti $ex = xe = x$ per ogni elemento x di G;
c) ogni elemento di G possiede un inverso rispetto a e, ossia un elemento x' di G tale che $xx' = x'x = e$.
Esempi di gruppi sono i seguenti: l'insieme Z dei numeri interi quando si assuma

Fatto questo breve cenno al contributo di Ruffini, occupiamoci dei due matematici sopra citati. Niels Henrik Abel nacque a Findö in Norvegia nel 1802; di ingegno matematico precocissimo, trova nel proprio professore Berndt Michael Holmboë (1795-1850), da lui conosciuto nel 1818, incoraggiamenti e aiuti finanziari che in specie dopo la morte del padre, avvenuta nel 1820, gli permettono di mantenere faticosamente la numerosa famiglia che ora grava tutta sulle sue spalle. Nel 1821 ritiene di aver scoperto un metodo per la risoluzione algebrica delle equazioni di grado superiore al quarto; invia la dimostrazione a un matematico danese, ma prima ancora che questi gli risponda si accorge di aver commesso un errore; affronta quindi il problema dal punto di vista opposto e riesce a dimostrare l'*impossibilità* di tale soluzione nella fondamentale *Mémoire sur les équations algébriques où on démontre l'impossibilité de la résolution de l'équation générale du cinquième degré* (*Memoria sulle equazioni algebriche dove si dimostra l'impossibilità della risoluzione dell'equazione generale di quinto grado*) che pubblica nel 1824 (e due anni dopo sul «Journal de Crelle»). Abel vi dimostra sostanzial-

come legge di composizione interna l'ordinaria operazione di addizione (elemento neutro il numero 0, inverso di un numero intero a il numero $-a$); l'insieme di tutti i numeri razionali positivi (le frazioni) quando per legge di composizione si assuma l'ordinaria operazione di moltiplicazione tra frazioni (elemento neutro la frazione 1/1, elemento inverso di una frazione a/b la frazione b/a). Esempi fondamentali di gruppi si presentano anche in geometria, come abbiamo avuto occasione di ricordare parlando di Klein e di Lie. Limitandoci al piano, consideriamo ad esempio il gruppo dei movimenti rigidi (traslazioni, rotazioni, simmetrie) in cui l'elemento neutro è la trasformazione identica e ove la legge di composizione interna è semplicemente l'applicazione successiva di trasformazioni. Altro gruppo è quello delle similitudini, altri quello delle affinità, delle proiezioni, quello delle trasformazioni continue. Per quanto riguarda la teoria delle equazioni il concetto di gruppo si presenta parlando delle sostituzioni o permutazioni. Consideriamo un insieme finito, fissiamo per comodità la nostra attenzione ad esempio sull'insieme $\{1, 2, 3\}$ costituito dai primi tre numeri naturali. Possiamo intendere come sostituzione o permutazione degli elementi di questo insieme un'operazione che altera in un modo qualunque l'ordine di questi elementi, al limite lasciandolo invariato o, nel caso di un insieme qualunque, una funzione biiettiva. Valendoci di una notazione ormai usuale indichiamo come segue le sostituzioni possibili nel caso particolare dell'insieme considerato nel nostro esempio (si può dimostrare in generale che se un insieme è composto da n elementi, sono possibili $n! = n(n-1)(n-2)\ldots 3 \cdot 2 \cdot 1$ permutazioni)

$$\left\{\begin{pmatrix}1\,2\,3\\1\,2\,3\end{pmatrix},\begin{pmatrix}1\,2\,3\\1\,3\,2\end{pmatrix},\begin{pmatrix}1\,2\,3\\2\,1\,3\end{pmatrix},\begin{pmatrix}1\,2\,3\\2\,3\,1\end{pmatrix},\begin{pmatrix}1\,2\,3\\3\,1\,2\end{pmatrix},\begin{pmatrix}1\,2\,3\\3\,2\,1\end{pmatrix}\right\}.$$

mente che le radici di un'equazione generale di quinto grado non possono essere espresse in funzione dei coefficienti dell'equazione per mezzo di radicali. La memoria venne anche inviata dal giovane matematico norvegese a Gauss, che però non le prestò la sia pur minima attenzione.

Nel 1825, ancora dietro interessamento di Holmboë, Abel riceve una modesta sovvenzione dal governo scandinavo per compiere un viaggio di studio in Germania e in Francia. Se pure il contatto con i grandi matematici continentali, vecchi e giovani (egli conobbe ad esempio Cauchy, Laplace, Lacroix) non gli fu di particolare profitto anche e soprattutto per l'atteggiamento di sufficienza che in generale questi assunsero nei suoi riguardi, egli tuttavia incontra a Berlino l'ingegnere August Leopold Crelle (1780-1856) che nel 1826 fonda il suo «Journal für die reine und angewandte Mathematik» («Giornale di matematica pura e applicata», noto anche semplicemente come «Journal de Crelle»). Intuita la genialità del giovane matematico Crelle pubblicherà sulla sua rivista gran parte della produzione di questi. Abel rientra in patria da questo viaggio nel 1827, è già gravemente ammalato di tisi e muore nel 1829. Due giorni dopo la sua morte giunge una lettera di Crelle con l'annuncio che l'università di Berlino lo chiamava per affidargli una cattedra di matematica.

Ma lo scienziato che, oltre a mostrare come Ruffini e Abel la non risolubilità per radicali delle equazioni algebriche generali di grado

Consideriamo ora l'insieme i cui elementi sono tutte e sole queste sostituzioni e conveniamo di definire fra gli elementi di questo nuovo insieme un'operazione (che chiameremo «prodotto» di sostituzioni) che consiste semplicemente nell'applicare successivamente due date sostituzioni; così ad esempio il prodotto della seconda e della quarta sostituzione sopra scritte sarà

$$\begin{pmatrix} 1\,2\,3 \\ 1\,3\,2 \end{pmatrix} \cdot \begin{pmatrix} 1\,2\,3 \\ 2\,3\,1 \end{pmatrix} = \begin{pmatrix} 1\,2\,3 \\ 2\,1\,3 \end{pmatrix}$$

e si può osservare quanto segue: il risultato dell'operazione così definita su una coppia qualsiasi di sostituzioni dell'insieme dato è sempre una sostituzione dello stesso insieme; tale operazione si può dimostrare essere associativa; la prima sostituzione sopra scritta, detta anche sostituzione identica, agisce come elemento neutro rispetto all'operazione fra sostituzioni (nel nostro esempio il lettore può verificarlo facilmente); ogni sostituzione del nostro insieme possiede una sostituzione inversa rispetto alla sostituzione identica; ad esempio, l'inversa della sostituzione

superiore al quarto, pone nel contempo in modo consapevole le basi di una nuova e fecondissima teoria è il francese Evariste Galois. Nato a Bourg-la-Reine nel 1811, muore dopo una vita tumultuosa e di profondo impegno politico a soli 21 anni, in seguito a un duello, nel 1832. I suoi scritti rimasero in lunga parte non pubblicati ed una presentazione complessiva dei suoi lavori si trova in alcune pagine scritte la notte precedente il duello nel quale doveva trovare la morte. Esse sono indirizzate all'amico A. Chevalier con una lettera nella quale, fra l'altro, lo prega di «chiedere a Jacobi e Gauss la loro opinione sulla importanza, non sulla verità, dei teoremi». «Dopo ciò,» proseguiva, «spero vi sarà qualcuno che troverà vantaggioso decifrare tutti questi sgorbi (*ce gâchis*).»

Galois ha una concezione completa della teoria dei gruppi di sostituzione e la sua scoperta essenziale riguardo alla teoria delle equazioni algebriche consiste nel mostrare come a ogni equazione algebrica sia associato un gruppo di sostituzioni delle radici (il «gruppo dell'equazione») nel quale sono riflesse le sue caratteristiche essenziali rispetto alla risolubilità. Ovviamente molti dei risultati di Galois erano indicati nelle sue carte senza dimostrazione; ma da quelle carte prende lo spunto la moderna teoria dei gruppi astratti: Joseph Liouville (1809-82) pubblica nel 1846 nel suo «Journal de mathématiques pures et appliquées» la maggior parte delle carte di Galois; la comprensione completa delle idee di questo giovane genio della matematica e la fecondità delle stesse si ebbe però solo a partire dal 1870, dopo cioè la pubblicazione del *Traité des substitutions* (*Trattato sulle sostituzioni*) di Camille Jordan (1838-1922).

$$\begin{pmatrix} 1\,2\,3 \\ 3\,1\,2 \end{pmatrix}$$

è la sostituzione

$$\begin{pmatrix} 1\,2\,3 \\ 2\,3\,1 \end{pmatrix}$$

dal momento che il prodotto di queste due sostituzioni è proprio uguale (lo si verifica ancora facilmente) alla sostituzione identica. Per quanto sopra detto, si può allora concludere che le sostituzioni sull'insieme finito $\{1, 2, 3\}$ formano un gruppo (e la cosa si può generalizzare a un insieme finito qualsiasi). Si ricordi ora che per il teorema fondamentale dell'algebra un'equazione algebrica di grado n ammette al più n radici distinte, per comprendere come abbia senso parlare di gruppo delle sostituzioni delle radici di una data equazione. Ricordiamo infine che per sottogruppo di un dato gruppo G si intende un qualunque sottoinsieme G' di G che risulti essere un gruppo rispetto all'operazione gruppale di G.

La novità maggiore di Galois rispetto agli studiosi continentali della sua epoca sta proprio nell'aver capito l'importanza dei gruppi di trasformazioni e del loro intergioco con il processo di estensione dei domini di razionalità (quelli che oggi chiamiamo campi). Galois sottolinea in modo netto il carattere astratto dei suoi risultati e la loro validità generale per campi arbitrari ed è proprio a partire da lui che, attraverso un lungo e faticoso processo, l'algebra si trasformerà da studio di equazioni algebriche (in definitiva quindi sempre ancorato all'aspetto *numerico* delle stesse) in studio generalissimo di *strutture algebriche*, come appunto sono ad esempio i gruppi ed i campi. La nuova impostazione permetterà fra l'altro di scoprire il significato più profondo di operazioni già note da tempo: la risolubilità o irresolubilità delle equazioni potrà così finalmente uscire dal campo della pratica aritmetica per entrare in quello più rigoroso della consapevolezza razionale, fornendo un esempio estremamente significativo di come affrontare in termini matematici problemi di carattere logico generale come appunto quello dell'esistenza di metodi per stabilire la risolubilità algebrica delle equazioni. Non è un caso che proprio alla teoria di Galois Hilbert faccia riferimento nel 1899 per illustrare il significato dell'indagine metamatematica dei sistemi assiomatici.

3.2 *L'algebra in Inghilterra*

L'orientamento implicito che sostanzialmente si può ritrovare dietro il nuovo atteggiamento della matematica continentale è quindi quello di giungere a dare una fondazione logica autonoma ai sistemi matematici, nel senso di scindere il problema della costruzione sistematica e strutturale da quello del riferimento a specifiche realtà e ad enti particolari (e ciò vale, pur se in misura diversa, tanto per l'Analisi quanto per l'algebra o la geometria).

Che questo orientamento sia comune a tutto il primo Ottocento europeo si vede in misura ancora maggiore ed esplicita dalle vicende legate alla scuola analitica di Cambridge in Inghilterra, dove una ristretta compagine di ricercatori, pur partiti in netto svantaggio rispetto alla mole di risultati dei continentali, riuscì ad accogliere questi risultati, sviluppando tuttavia nel contempo una tematica originale soprattutto per quanto riguarda il chiarimento

della natura dell'algebra, e il progressivo distacco dell'algebra astratta o simbolica da quella concreta, legata ancora cioè al momento numerico. Questo processo troverà in effetti una prima esplicita e consapevole formulazione proprio nei lavori logici di George Boole di cui, come già avvertito, ci occuperemo nel prossimo capitolo. Nelle pagine seguenti illustreremo brevemente la problematica della scuola inglese, dalla quale doveva poi scaturire in tutta pienezza il concetto di ragionamento formale, «attraverso un processo che porta all'analisi delle strutture formali dell'algebra sino all'algebra della logica» (Barone) e che quindi per noi è particolarmente interessante.

Questa problematica si radica e prende spunto dall'isolamento in cui l'ambiente scientifico inglese, profondamente newtoniano e quindi particolarmente legato in Analisi al metodo e alla notazione flussionali, era venuto a trovarsi rispetto all'enorme massa di risultati che scienziati come Eulero, Lagrange, Laplace avevano fatto accumulare, nel continente, all'Analisi matematica, grazie anche all'impiego del metodo e della notazione differenziali leibniziani. Si ricordi infatti che la prima edizione completa dell'*opera omnia* di Newton era stata pubblicata in Inghilterra fra il 1776 e il 1785 a cura di Samuel Horsley (1733-1806) e, ad esempio, che nella traduzione inglese (1801) delle *Istituzioni analitiche* di Maria Gaetana Agnesi (1718-99) tutta la notazione leibniziana in termini differenziali era stata sostituita in termini flussionali.

Già tuttavia in questo periodo si avverte la debolezza di tale posizione e si hanno i prodromi di un vigoroso e vitale movimento che intende por fine a quel «ridicolo isolamento» della scuola matematica inglese che doveva preparare a sviluppi più strettamente attinenti alla nostra storia. Iniziatore di tale movimento può considerarsi Robert Woodhouse (1774-1827) che pubblica a Cambridge nel 1803 il volume *Principles of analytical calculations* (*Principi di calcoli analitici*) nel quale critica, talora con caustico sarcasmo, l'esclusivo impiego di metodi e notazioni flussionali e nel quale tuttavia assume anche un originale e corretto atteggiamento critico nei riguardi dei risultati degli analisti continentali, motivato esplicitamente dall'assenza di rigore con la quale essi operavano. Questo volume avrà alla lunga una notevole influenza nell'ambiente matematico di Cambridge e indubbiamente ebbe non piccola parte nella decisione, presa da alcuni giovani ricercatori di

tale università, di costituire nel 1812 una Società analitica che si assunse il compito di indire riunioni periodiche per la diffusione della notazione differenziale e per la condanna della cosiddetta «eresia dei puntini».

La Società venne costituita, fra gli altri, ad opera di Charles Babbage (1792-1881), George Peacock (1791-1858) e John William Herschel (1792-1871), figlio del celebre astronomo e astronomo a sua volta. La Società si propone di porre «i matematici inglesi sullo stesso piano dei loro rivali continentali» e si propone anche la traduzione dei tre volumi del *Traité de calcul différentiel et intégral* (*Trattato di calcolo differenziale e integrale*) del francese François Lacroix (1765-1843) pubblicati a Parigi fra il 1797 e il 1800, traduzione che venne eseguita dai tre scienziati sopra nominati e che venne pubblicata a Cambridge nel 1816 e seguita da due volumi di esempi e esercizi. A cura della Società esce nel frattempo, curato da Babbage e da Herschel, un volume di *Transactions* dal titolo *The principles of pure D-ism in opposition to the Dot-age of the university* (*I principi del puro D-ismo contrapposti all'Età-del-puntino dell'università*). La conquista graduale del nuovo metodo era così cominciata e la si può considerare completa in Inghilterra attorno al 1830.

Ma tutto questo fervore non era fine a se stesso: non serviva soltanto a rompere lo splendido isolamento britannico. Ci sembra che il momento di maggior rilievo di questo processo più positivo e significativo venga colto molto bene da Francesco Barone quando afferma che, tramite esso «ci si avviò ad abbandonare l'idea che la validità di una dimostrazione dipenda solo dalla natura specifica dell'argomento trattato, a considerare la matematica pura in cui contano le proprietà formali comuni ad argomenti diversi, a non valersi più esclusivamente di interpretazioni geometriche come fondamento dei concetti matematici ma, anzi, a guardare la stessa geometria pura nel suo aspetto assiomatico quale sistema linguistico di simboli. Furono questi atteggiamenti che agirono in profondità e condizionarono, al momento opportuno, l'atteggiamento dei membri della scuola di Cambridge di fronte al problema dei fondamenti della matematica». Ove va ricordato che era una «caratteristica propria della tradizione flussionale di trattare i problemi tutti mediante la geometria e di dare alle dimostrazioni geometriche il predominio esclusivo».

Ora però può riscontrarsi in questa attenzione ai problemi fondazionali che accomuna nel primo Ottocento l'Europa matematica con l'Inghilterra, una differenza sottile che ci sembra tuttavia importante e proficuo mettere in luce. L'atteggiamento continentale in questo periodo è strettamente legato, si può dire è frutto diretto, e in certo senso conseguente, dell'eredità analitica del Settecento: l'esigenza di rigore è per così dire «intrinseca», è interna a un sistema costituito cui si vogliono dare basi più solide assicurando una fondazione rigorosa ai concetti di numero reale, di infinitesimo, di limite. Si vuole in altri termini dare una soddisfacente e sicura fondazione alle basi per così dire «concrete» di un edificio già tanto glorioso e monumentale.

L'aver gli inglesi ricevuto questi risultati senza una «tradizione» prossima di copiosa ricerca analitica alle spalle, li pone in un certo senso in una condizione di favore, ossia permette loro, invece di dedicarsi alla sistemazione di tale edificio, di valutare globalmente la matematica intesa come «scienza delle quantità», «teoria delle grandezze» giusta la definizione di numero (ancora esplicitamente espressa, ad esempio, da Eulero) quale rapporto fra grandezze. È ovvio che qui la prospettiva è, al contempo, più ambiziosa e difficile: da una parte è chiaro infatti che in questo modo si presagiva una teoria più generale delle forme, proprio in senso leibniziano, di cui la matematica in quanto applicata a grandezze geometriche o aritmetiche non sarebbe stata che un caso particolare; d'altra parte era così radicato il legame fra operazioni algebrico-analitiche in senso lato e enti – per l'appunto numerici o geometrici – cui queste operazioni si applicavano, che ben arduo si presentava il compito di scindere la struttura algebrica dalla natura specifica delle operazioni in esame raggiungendo così una visione astratta della matematica.

I matematici inglesi si applicarono in questo senso con particolare riguardo all'algebra, intesa fino ad allora come una sorta di generalizzazione dell'aritmetica. Un ostacolo non trascurabile che si opponeva di principio ad una concezione astratta dell'algebra era proprio rappresentato dal modo stesso di intendere il concetto di numero ereditato dal Settecento: assunto come «dato» – su basi non logicamente motivate – il numero naturale, si giustificava la liceità dell'impiego dei numeri relativi, immaginari, ecc. fondandosi su modelli *di fatto* desunti dalla loro esprimibilità come rap-

porto di grandezze, in particolare geometriche (si ricordi tuttavia che ancora Eulero giustificava l'introduzione dei numeri relativi col modello dei crediti e dei debiti).

Malgrado non fossero mancati precedenti (ad esempio Clairaut, 1794, e Laplace, 1812) che tentavano di giustificare l'introduzione di nuovi numeri ricorrendo all'esame delle proprietà formali delle operazioni con gli stessi, va anche detto che proprio in questo periodo sembra venire una conferma contraria alla necessità di giustificare su basi logiche le nuove entità numeriche piuttosto che limitarsi ad accettarle pragmaticamente sulla base di un modello di un qualche tipo. Il vecchio problema di giustificare l'intervento nei calcoli delle mal tollerate quantità immaginarie viene infatti risolto in questo periodo a partire dalla memoria di Caspar Wessel del 1797, dai lavori di Jean Robert Argand del 1813 e fino alla definitiva consacrazione nel 1831 da parte di Gauss, proprio sulla base del noto modello geometrico (il piano di Argand-Gauss) per cui i numeri complessi possono essere interpretati come «punti» di un piano riferito a un sistema cartesiano i cui assi rappresentino l'uno le quantità «reali», l'altro le quantità «immaginarie» che figurano nell'espressione generale $a + ib$ di un numero complesso (a, b, reali, $i = \sqrt{-1}$, unità immaginaria). Se ciò da una parte pareva dimostrare che le operazioni algebriche non sono connesse in modo essenziale e indissolubile ai «principi e alle definizioni delle quantità ordinariamente considerate nell'aritmetica», segnando un distacco dell'algebra dall'aritmetica, ne sembrava per altro verso discendere in modo addirittura conclusivo un suo stretto ed essenziale legame con la geometria, che però è all'insegna di una prospettiva diversa da quello tradizionale: il sogno leibniziano di un calcolo geometrico che considera *direttamente* le figure e i loro rapporti in veste algebrica senza passare attraverso coordinate numeriche. Non va dimenticato infatti che il problema di Wessel e in parte di Argand non era quello di giustificare le regole del calcolo con i numeri complessi, quanto quello di elaborare una teoria delle grandezze orientate nello spazio. Anche qui il ricorso al formalismo, il distacco da interpretazioni tradizionali, non è fine a se stesso ma nasce dal desiderio di ampliare le potenzialità dei metodi analitici mantenendo la freschezza dell'approccio geometrico.

Tocca ancora a Woodhouse essere l'iniziatore *ante litteram*, in

Inghilterra, della reazione a ogni concezione preconcetta dei rapporti fra algebra e geometria: è del 1801 infatti un suo articolo *On the necessary truth of certain conclusions obtained by means of imaginary quantities* (*Sulla verità necessaria di certe conclusioni ottenute mediante quantità immaginarie*) nel quale egli pone, per così dire, il problema «logico» della giustificazione degli immaginari ascrivendo le «stranezze» e le «difficoltà» cui essi danno luogo proprio alla mancanza di una corretta spiegazione «della loro natura e del loro impiego» e sostenendo quindi che le quantità immaginarie «poiché portano a conclusioni esatte, devono pur avere una logica». È qui già avvertibile la tendenza al disancoramento dell'algebra tanto dall'aritmetica quanto dalla geometria, da quelle interpretazioni privilegiate che a quel tempo sono intese come elementi intrinsecamente costitutivi dell'algebra stessa. Questo nuovo atteggiamento si desume ancor più chiaramente da un secondo articolo dello stesso Woodhouse pubblicato nell'anno successivo e intitolato *On the independence of the analytical and geometrical methods of investigation* (*Sull'indipendenza dei metodi di ricerca geometrico e analitico*) ove l'autore afferma esplicitamente che «per i sistemi dell'analisi e dell'algebra le espressioni di origine geometrica non hanno posizione privilegiata e il loro impiego è spesso legato soltanto alla convenienza grafica dei loro simboli» (ove fra l'altro si avverte la precorritrice dimensione linguistica della ricerca astratta).

L'influenza di queste prese di posizione di Woodhouse fu molto feconda anche se non immediatamente evidente, e venne maturando in concomitanza con l'affermarsi anche in Inghilterra di un principio, assai diffuso fra i matematici del tempo, detto «della separazione dei simboli» cui aveva dato sistemazione organica nel 1814 il francese François-Joseph Servois (1767-1847) col suo articolo *Essai sur un noveau mode d'exposition des principes de calcul différentiel* (*Saggio su un nuovo modo di presentare i principi di calcolo differenziale*). Esso consisteva sostanzialmente nel non ritenere legati (nel «separare», appunto) i risultati dell'applicazione di certe operazioni analitiche (ad esempio, la derivazione) con gli argomenti cui esse erano applicate, cercando di stabilire un'«algebra» per questi simboli, algebra che valesse per loro in quanto tali, indipendentemente dagli argomenti cui avrebbero potuto essere applicati (considerandoli cioè, si potrebbe dire in linguaggio moderno, co-

me una sorta di operatori del tutto generali).[19] L'articolo di Servois trovò eco in Inghilterra in un lavoro di Herschel comparso nello stesso anno, *Considerations of various points of analysis* (*Considerazioni su vari argomenti in Analisi*), e anche un successivo lavoro dello stesso autore, *Collection of examples on the calculus of finite differences* (*Raccolta di esercizi sul calcolo delle differenze finite*) del 1820, si ispirava allo stesso metodo.

Le idee di Woodhouse vengono riprese in particolare da Peacock, il quale inoltre si serve del principio suddetto applicandolo però all'algebra. Nel 1834 egli compila un *Report on the recent progress and the present state of certain branches of analysis* (*Relazione sui recenti progressi e sullo stato attuale di alcune branche dell'Analisi*), ove ribadisce la separazione fra la teoria generale dell'algebra da un punto di vista simbolico (astratto) e le sue interpretazioni in termini geometrici o di altro tipo. In un contesto più generale, Peacock realizza chiaramente e afferma esplicitamente la distinzione fra scienze astratte («speculative») e applicate («fisiche») con un'argomentazione di cui conviene riportare qualche passo più significativo. «Nelle scienze speculative», afferma Peacock, «consideriamo soltanto i risultati della scienza stessa e il rigore logico del ragionamento con cui essi vengono dedotti dai primi principi assunti e tutte le nostre conclusioni posseggono la necessaria esistenza indipendentemente dalla interpretazione più o meno aderente della natura delle cose. Nelle scienze fisiche i nostri ragionamenti

[19] Ecco ad esempio come Gregory (citato più avanti nel testo) chiarisce questo principio: «Nell'algebra ordinaria si ha un certo numero di teoremi che, malgrado sembrino dimostrati soltanto per i simboli che rappresentano numeri, ammettono un'applicazione più estesa. Tali teoremi dipendono soltanto dalle leggi di combinazione a cui i simboli sono sottoposti, e quindi sono veri per tutti i simboli, quale che sia la loro natura, che sono sottoposti alle stesse leggi di combinazione. Le leggi di cui trattiamo sono poche e possono essere espresse come segue: siano a e b due operazioni e u e v due loro possibili argomenti; allora le leggi sono: 1) $ab(u) = ba(u)$; 2) $a(u + v) = a(u) + a(v)$; $a^m a^n(u) = a^{m+n}(u)$. La prima di queste leggi è chiamata *commutativa*, la seconda *distributiva*,... Che queste siano le leggi usate nella dimostrazione dei principali teoremi in algebra può essere facilmente mostrato da un breve esame dei procedimenti; ma esse non sono limitate ai simboli dei numeri: si applicano anche al simbolo usato per denotare la differenziazione. Infatti se u è una funzione di due variabili x e y, da noti teoremi del calcolo differenziale si ha

$$\frac{d}{dx}\frac{d}{dy}(u) = \frac{d}{dy}\frac{d}{dx}(u)\ldots».$$

si basano ancora sui primi principi assunti, e analogamente osserviamo l'accuratezza logica delle deduzioni; ma tanto nei principi quanto nelle conclusioni dedotte guardiamo al mondo esterno che ci fornisce attraverso l'interpretazione principi e conclusioni corrispondenti... Pertanto i principi primi... nelle scienze fisiche, non essendo assunzioni arbitrarie né verità necessarie, ma facendo parte delle proposizioni che costituiscono la scienza, non possono mai cessare di essere indagati ed esaminati in qualche punto delle nostre ricerche... Ma nelle scienze astratte come la geometria e l'algebra i principi che ne costituiscono il fondamento rappresentano anche il limite proprio della nostra ricerca, perché, se in qualche modo quelle scienze speculative sono connesse con le scienze fisiche, la connessione è arbitraria e non tocca la verità delle nostre conclusioni, che riguarda soltanto la connessione coi principi primi e non richiede, benché permetta, l'aiuto dell'interpretazione fisica».

Peacock aveva fatto anche un tentativo di concreta applicazione di queste idee, pubblicando nel 1830 *A Treatise on algebra* (*Trattato di algebra*) ove presenta la materia da un punto di vista «speculativo» o «astratto». In una seconda edizione di questa opera in due volumi, rispettivamente del 1842 e del 1845, egli dedica il secondo volume all'algebra simbolica, ribadendo il fondamentale principio che il «significato» delle operazioni e dei risultati dipende propriamente soltanto dai postulati assunti e non dalle interpretazioni dei simboli; il che sembra esprimere sostanzialmente, anche se in modo non esplicito, l'idea che un sistema ipotetico-deduttivo formale è suscettibile di *diverse* interpretazioni, così che il momento semantico è essenzialmente distinto dalla costituzione sintattica del sistema che ne è, in linea di principio, indipendente.

Ma questo passo non viene compiuto da Peacock: il germe deve ancora aspettare per dare i suoi frutti. In effetti l'analisi peacockiana vede pur sempre la «fondazione» dell'algebra astratta nell'algebra aritmetica, non più considerata come interpretazione, ma come «suggeritrice» di forme affatto generali: nello stesso *Report* Peacock sostiene infatti che «disancorandosi completamente da essa non saremmo del tutto in grado di interpretare sia le nostre operazioni sia i nostri risultati, e la scienza che ne risulterebbe sarebbe costituita da tali simboli che non ammetterebbero interpretazioni di nessun tipo». Sembra così di intravvedere sullo sfon-

do un problema logico ben preciso: entro che limiti, se diamo regole (o assiomi *formali*) senza avere in mente un'interpretazione determinata, possiamo essere sicuri che le regole non saranno contraddittorie (o che almeno un'interpretazione di esse esista)? Vedremo più avanti come Boole affronti il problema che costituirà il banco di prova di ogni concezione di tipo formale o simbolico una volta che – abbandonato l'aggancio a dati concreti – ci si trovi di fronte alla questione di come stabilire a priori criteri di validità e di interpretabilità dei sistemi di regole. Quello a cui Peacock giunge è un «principio di permanenza delle forme equivalenti» che rimanda essenzialmente al modello concreto dell'algebra sui numeri reali e vede il costituirsi del ragionamento simbolico come estensione di queste stesse regole a domini più ampi di oggetti mantenendo così la stessa struttura formale e lasciando eventualmente cadere limitazioni riguardo all'effettuabilità delle operazioni (ad esempio divisione, sottrazione, ecc.).

Sostanzialmente aderente alle conclusioni di Peacock è anche D.F. Gregory (1813-44), secondo il quale «... il passo fatto dall'algebra aritmetica a quella simbolica consiste in ciò che, non considerando la natura delle operazioni che i simboli rappresentano, supponiamo l'esistenza di classi di operazioni sconosciute soggette alle stesse leggi. Possiamo così dimostrare l'esistenza di certe relazioni, e queste relazioni, se espresse fra simboli, vengono chiamate teoremi algebrici». Se pure in Gregory sembra intravvedersi una maggior apertura verso la possibilità di interpretazioni qualsiasi, la sua posizione si adegua essenzialmente al principio di permanenza di Peacock.

Molto opportunamente Barone collega questo principio peacockiano (che può essere considerato il punto d'arrivo dell'elaborazione della scuola di Cambridge in questa direzione) con un altro celebre principio, detto di «permanenza delle leggi formali», che il matematico tedesco Hermann Hankel (1839-1873) enuncerà nel 1867. Questo ultimo principio è stato sottoposto a severe critiche dagli algebristi moderni proprio per il residuo ancoramento a interpretazioni privilegiate che esso esprime, e che in definitiva impedisce di cogliere in tutta la sua portata la piena libertà della costruzione formale. Le stesse critiche abbiamo in effetti mosso al principio di Peacock, ma ci sembra di poter aderire alla tesi di Barone convenendo che «... nonostante il complesso di pro-

blemi lasciato aperto... dalla realizzazione ottocentesca della concezione di una matematica pura, sarebbe storicamente ingiusto, specie nei confronti di Peacock, non riconoscere... l'importanza di ciò che essa ha attuato... Senza l'insistenza sulle proprietà formali delle consuete operazioni aritmetiche l'attenzione non si sarebbe appuntata sui caratteri del procedimento formale in genere».

Nell'ambiente algebrico inglese così delineato si inserisce di diritto il matematico irlandese William Rowan Hamilton[20] il quale, pur non appartenendo alla scuola di algebristi di Cambridge, era in relazione con alcuni esponenti della scuola stessa, ad esempio con Peacock e con De Morgan e su questi problemi assume posizioni diverse.

Hamilton individua tre scuole con posizioni sostanzialmente differenti «... la pratica, la filologica e la teoretica, a seconda che l'algebra stessa sia considerata uno strumento, o un linguaggio, o una contemplazione». In questa sua terminologia la scuola filologica era appunto quella di Cambridge; ed egli non concorda con l'analisi che dell'algebra aveva condotto questa scuola, in particolare con la conclusione cui i suoi rappresentanti erano giunti, sulla possibilità di sviluppare un'algebra puramente simbolica (abbiamo sopra visto, peraltro, con quale riserva e limitazione vada assunta questa conclusione). Hamilton non riesce a concepire l'idea di un'algebra completamente astratta, e si rifiuta di attribuire «l'alto nome di scienza» a un «sistema di simboli e niente più, a una questione di ganci e uncini, di segni neri su carta bianca, da farsi secondo una serie di regole stabilite ma arbitrarie... È una cortesia esagerata parlare del giuoco degli scacchi come di una "scienza" benché esso possa essere chiamato un gioco scientifico».

Mentre concorda con gli algebristi di Cambridge circa l'esigenza di non più considerare la matematica come «teoria delle gran-

[20] Nacque a Dublino nel 1805. Di intelligenza estremamente precoce, entrò nel 1823 nel Trinity College di Dublino e solo quattro anni dopo, nel 1827, venne nominato professore di astronomia e astronomo reale all'Osservatorio di Dunsink. Occupatosi dapprima intensamente di ottica geometrica e dinamica, dopo che nel 1843 ebbe scoperto il sistema dei quaternioni si dedicò quasi esclusivamente al suo perfezionamento e alla ricerca di sue possibili applicazioni. Morì a Dublino nel 1865. Oltre alle opere citate nel testo ricordiamo i postumi *Elements of quaternions* (*Elementi dei quaternioni*) del 1866.

dezze», Hamilton ritiene quindi necessario non fermarsi alla esclusiva considerazione «linguistica» e simbolica dell'algebra, ma cercare il «puro» contenuto denotato dai simboli algebrici. In parte – ma solo ad uno stadio avanzato della ricerca – riferendosi all'estetica trascendentale kantiana, egli crede di poter individuare questo contenuto nel *tempo*, inteso come possibilità di ordinamento in continua progressione. Come scrive Hamilton, l'algebrista teorico non vede nell'algebra un puro strumento né un semplice linguaggio in cui *formulare* risultati che non è in grado di *dimostrare* algebricamente e «si lamenta dell'imperfezione, di come viene oscurata la sua contemplazione, quando i ragionamenti della sua scienza sembrano ovunque opporsi l'uno all'altro e divenire o troppo complessi o troppo poco validi per potercisi fondare sopra o quando, anche se la verifica ha dimostrato che una regola è utile o che una formula dà risultati veri, non può provare questa regola né comprendere la formula: quando non può sollevarsi all'intuizione dall'induzione o non può volgere gli occhi oltre i segni alle cose significate». Quello che Hamilton cerca per l'algebra (che – come sottolinea accuratamente – non va confusa con l'Analisi generale di cui parlava la scuola di Cambridge) è da una parte la determinazione di un contenuto reale, di un modello (il che non significa però accettare l'algebra come teoria delle grandezze) dall'altra una fondazione che sia in grado di giustificare razionalmente, con riferimenti ad un ben preciso tipo di intuizioni, la validità e l'applicabilità delle regole algebriche.

Non deve stupire quindi che il massimo contributo di Hamilton al dibattito sulla natura dell'algebra abbia avuto come punto focale la costruzione di *modelli*. Fu questo che lo spinse a costruire sistemi che rompevano con leggi tradizionalmente inattaccabili come la commutatività. Non è un caso che anche algebristi della scuola filologica come De Morgan fossero turbati da questi sviluppi, che venivano ad infrangere l'unico principio – seppur vago – che nella loro concezione giustificava il ragionamento simbolico: la permanenza delle leggi formali rispetto all'algebra tradizionale. In un senso, se ci si manteneva nell'orizzonte della concezione formale, non c'era alternativa per garantirsi che «i ragionamenti della scienza non si contraddicessero l'uno con l'altro». L'idea di giustificare nuovi sistemi algebrici ricorrendo a principi di transfer (come diremmo noi oggi) non aveva modo di spiegare l'introdu-

zione di regole che contraddicevano quelle valide nei domini di partenza. Questo era possibile costruendo modelli delle nuove regole a partire da modelli per le vecchie. Il metodo di Hamilton si basa essenzialmente sulla considerazione di coppie (come nel caso dei negativi, dei razionali o dei complessi) di numeri noti o di triple o quadruple (come nel caso dei quaternioni) e sulla definizione di operazioni su di esse a partire da quelle sui domini noti. È difficile sottovalutare l'importanza di questo passo, che si muove nella direzione delle interpretazioni piuttosto che delle forme e che in certo senso è complementare alle prospettive aperte dalla scuola filologica. Hamilton mostra come dato un calcolo sia possibile fornirne una semantica, un'interpretazione. Naturalmente per Hamilton il problema non è un esercizio logico; egli mira ad un calcolo geometrico per le figure orientate in cui il verso, la progressione, l'ordine abbiano un posto e non ci si limiti alle semplici grandezze. È qui che interviene l'idea di *tempo*, come fonte della distinzione fra prima e dopo, che è alla base della considerazione di coppie, di triple, di *n*-uple. Il metodo è generale e per Hamilton mostra come il vero oggetto dell'algebra siano le quantità variabili e l'algebra sia la scienza delle funzioni. È su questo terreno che seppure con una consapevolezza minore delle potenzialità del ragionamento formale, le ricerche di Hamilton si avvicinano alle contemporanee ricerche di Hermann Grassmann (1809-1877).[21]

La prima esposizione sistematica di queste idee si trova nel saggio del 1833, *Theory of coniugate functions or algebraic couples; with a preliminary and elementary essay on algebra as the science of pure time* (*Teoria delle funzioni coniugate o coppie algebriche, preceduta da un saggio elementare sull'algebra come scienza del tempo puro*) che fornisce la giustificazione «logica» (la prima che sia mai stata data) dei numeri complessi. Nella prima parte del saggio Hamilton giunge a costituire il campo dei numeri reali a partire dai numeri naturali presupposti come dati; nella seconda parte giustifica l'introduzione dei numeri complessi sulla base di un calcolo fra «coppie» di numeri reali – ossia di elementi della forma (a, b), (c, d) con a, b, c, d

[21] È opportuno tuttavia ricordare che successivamente Hamilton – come può dedursi dal suo epistolario con Peacock e Graves – sembra essersi avvicinato moltissimo, anche sotto questo riguardo, alla concezione degli algebristi di Cambridge circa la liceità di una matematica meramente simbolica e astratta.

reali – calcolo che gode delle stesse proprietà formali dell'usuale calcolo con i numeri complessi.[22]

Il secondo contributo è contenuto nelle *Lectures on quaternions* (*Lezioni sui quaternioni*) del 1853, dove, assieme alla presentazione di quella che l'autore ritiene la sua scoperta fondamentale – il sistema dei quaternioni, appunto – si ha la codificazione dell'algebra quale scienza del tempo puro; le *Lectures* erano state precedute nel 1844 dal saggio *On quaternions, or a new system of imaginaries in algebra* (*Sui quaternioni, ossia un nuovo sistema di immaginari in algebra*). Ancora, non interessa qui scendere in particolari circa questo nuovo sistema: ci basterà notare che si tratta di un sistema di enti esprimibili con quattro unità, una reale e tre immaginarie e che fra queste ultime, indicate con i, j, k, viene definita un'operazione di prodotto *che non gode della proprietà commutativa* (si ha ad esempio $ij = k, ji = -k$). Non sembra superfluo ribadire che il momento più interessante, nel nostro contesto, della scoperta hamiltoniana sta proprio in questo superamento *ante litteram* del principio di permanenza hankeliano con la conseguente costruzione di un sistema algebrico impostato a livello puramente formale. Si comprende anche meglio come William Rowan Hamilton, proprio di fronte a queste sue creazioni, fosse preoccupato di assicurarsi una «giustificazione» non arbitraria delle leggi che assumeva in questi nuovi sistemi, ricorrendo così all'interpretazione «temporale» che sostituiva alle grandezze «pure» le grandezze orientate.

È in questo ambiente culturale che si forma Boole e inizia la sua produzione nel 1840 con un lavoro sulle equazioni differenziali procedendo poi, dicevamo, nell'approfondimento in questa direzione. L'analisi di questo processo sarà oggetto di un prossimo capitolo; qui vogliamo tuttavia notare che l'occasione alla pubblicazione, da parte di Boole, del suo primo lavoro specificamente «logico» gli venne offerta, secondo quanto egli stesso dichiara, da una celebre disputa che a partire dal 1846 si svolse in Inghilterra fra Augustus de Morgan (cui avremo occasione di accennare par-

[22] Non è il caso di fermarci qui a illustrare i particolari di questa tecnica (che del resto avremo occasione di ricordare più avanti). Ci limiteremo a osservare che essa è tutt'oggi regolarmente impiegata nella presentazione logica dei sistemi numerici (sempre a partire dai naturali come dati, e una volta opportunamente definiti i numeri reali, ad esempio, come particolari successioni di numeri razionali).

lando di Boole) e il logico e filosofo inglese William Hamilton (da non confondersi con il matematico irlandese di cui abbiamo finora parlato), circa la priorità della scoperta dell'idea direttrice fondamentale di quello che può essere considerato l'ultimo tentativo di elaborazione tradizionale della sillogistica operato dallo stesso Hamilton. A questo punto è quindi opportuno prendere in esame proprio questi ultimi tentativi di sistemazione o di estensione della sillogistica tradizionale.

M. Kline, *Mathematical Thought from Ancient to Modern Times*, New York, Oxford University Press, 1972 [trad. it., *Storia del pensiero matematico*, 2 voll. (I. *Dall'antichità al Settecento*; II. *Dal Settecento a oggi*), Einaudi, Torino 1991].

M. Kline, *Mathematics in Western Culture*, Oxford University Press, 1953 [trad. it., *La matematica nella cultura occidentale*, Feltrinelli, Milano 1976].

U. Bottazzini, *Il flauto di Hilbert*, Utet Libreria, Torino 1990.

C.B. Boyer, *A history of Mathematics*, John Wiley & Sons, Inc., 1968 [trad. it., *Storia della matematica*, Isedi, Milano 1976].

D.J. Struik, *A concise history of Mathematics*, Dover Publications Inc., 1948 [trad. it., *Matematica: un profilo storico*, Il Mulino, Bologna 1981].

Euclide, *Elementi*, a cura di A. Frajese e L. Maccioni, Torino, Utet 1970.

R. Bonola, *La geometria non-euclidea. Esposizione storico-critica del suo sviluppo*, Zanichelli, Bologna 1906. Ristampa anastatica, 1975.

R. Trudeau, *The non-Euclidean revolution*, Birkhäuser, Boston 1987 [trad. it., *La rivoluzione non euclidea*, Bollati Boringhieri, Torino 1991].

E. Agazzi e D. Palladino, *Le geometrie non euclidee*, Mondadori, Milano 1978.

I. Toth, «Wann und von wem wurde die nichteuklidische Geometrie begründet?», in Archives Internat. Hist. Sciences, vol. 30/1980, pp. 193-205.

Id., «Un problème de logique et de linguistique concernant le rapport entre Géométrie euclidienne (GE) et Géométrie non-euclidienne (GNE)» in *Langage et Pensée Mathématiques* (Actes du Colloque International organisé au Centre Universitaire de Luxembourg les 9. 10 et 11 juin 1976, pp. 95-141).

Id., «Das Parallelenproblem im Corpus Aristotelicum» in *Archive for History of exact Sciences*, vol. 3, Number 4/5.

Id., «Mathematische Philosophie und hegelsche Dialektik. Ein Es-

say», in *Hegel und die Naturwissenschaften*, Band II der Reihe «Speculation und Erfahrung», Fromman-Holzboog, 1987, pp. 89-182.

Id., *La geometria non euclidea prima di Euclide*, in «Le Scienze», gennaio 1970, pp. 54-60.

J.J. Gray, *Ideas of space*, Clarendon Press, Oxford 1979.

R. Torretti, *Philosophy of Geometry from Riemann to Poincaré*, D. Reidel Publishing Co., Dordrecht 1978.

L. Novy, *Origins of Modern Algebra*, Noordhof International Publishing, Leida 1973.

B.L. Van der Waerden, *A History of Algebra*, Springer Verlag, Berlino 1985.

J.L. Richards, *Mathematical Visions*, Academic Press, San Diego 1988.

CAPITOLO SECONDO
IL RINNOVAMENTO ALGEBRICO DELLA LOGICA

1. LA LOGICA DEDUTTIVA NELLA PRIMA METÀ DELL'OTTOCENTO FINO A BOOLE

La cosa più sorprendente nel contesto degli studi logici della prima metà dell'Ottocento è che non si riscontrano – come forse sarebbe naturale attendersi – orientamenti e direttive di ricerca che in qualche modo lascino prevedere la svolta operata da Boole nel 1847 e nel suo complesso la ricerca logica si presenta come prosecuzione di quella settecentesca. Una nuova impostazione, caratteristica degli studiosi inglesi ma ravvisabile anche in autori continentali, relativa alla considerazione *estensionale* del calcolo sillogistico (logico), a nostro parere è di per sé troppo generica per permettere di ravvisare in essa (come peraltro alcuni autorevoli storici sostengono) un decisivo elemento di anticipazione dell'impostazione booleana della logica matematica. Non si vuole con ciò – ovviamente – privilegiare in alcun modo categorie come quella del «precorrimento» o della «anticipazione»; si intende solo sottolineare il fatto già espresso nelle pagine precedenti che altri fenomeni e ambiti culturali condizionarono positivamente il pensiero di Boole, almeno per quanto riguarda la costituzione del suo sistema, e in particolare appunto l'evoluzione della matematica dell'epoca.

Per quanto ora riguarda il continente segnaliamo intanto i due saggi *Réflexions sur la logique* (*Riflessioni sulla logica*) del 1802 e *Mémoire sur un nouvel algorithme logique* (*Memoria su un nuovo algoritmo logico*), presentati dal francese (di origine italiana) Gian Francesco Castillon nelle «Mémoires de l'Académie Royale de Sciences et Belles-Lettres de Berlin». Il secondo di questi due saggi è di gran lunga il più importante e in esso è sviluppata una versione intensionale della sillogistica. In questa sua sistemazione Castillon in-

troduce, accanto ai simboli A, S,... indicanti concetti (in particolare: soggetti e predicati di proposizioni categoriche) presi in intensione, una opportuna «indeterminata» M (che indica ovviamente ancora un concetto in intensione, ma lasciato appunto indeterminato) mediante la quale gli riesce di superare una difficoltà incontrata già da Leibniz e Lambert e relativa alla questione dell'inversione delle operazioni. Così, se $S + M$ rappresenta la «sintesi» di S e M,[1] e $S - M$ l'«astrazione» di M da S,[2] si ha la seguente simbolizzazione per le proposizioni aristoteliche:

A) Universale affermativa	Tutti gli S sono A	$S = A + M$
E) Universale negativa	Nessun S è A	$S = -A + M$

Le particolari corrispondenti vengono suddivise da Castillon in «reali» e «illusorie»;

I) Particolare affermativa	reale (è la conversa di una universale affermativa)	$A = S - M$
	illusoria (la sua conversa è a sua volta particolare)	$S = A \mp M$
O) Particolare negativa	reale	$A = -S + M$
	illusoria	$S = -A \mp M$

In questa sistemazione è immediatamente trascrivibile la teoria della conversione e della derivazione sillogistica. Una volta trascritte le premesse del sillogismo come sopra visto, si considerano i segni «+» e «−» come i segni aritmetici corrispondenti e si eseguono normali somme e riduzioni algebriche. Così ad esempio si ha

ogni M è A	$M = A + N$	nessun M è A	$M = -A + N$
ogni S è M	$S = M + P$	ogni S è M	$S = M + P$
ogni S è A	$S = A + (N + P)$	nessun S è A	$S = -A + (N + P)$

e così via.

[1] Ossia l'aggiunta di qualche ulteriore specificazione a S, sicché $S + M$ rappresenta una specie che contiene (intensionalmente) S.
[2] Ossia $S - M$ rappresenta un genere in cui è contenuto S (e M è la «differenza logica» di S in $S - M$).

Clarence Irving Lewis in *A survey of symbolic logic* (*Panorama della logica simbolica* del 1918, da cui abbiamo tratto la precedente illustrazione del sistema di Castillon) anche se considera quello di Castillon come «il miglior tentativo di istituire un calcolo logico in intensione» non manca tuttavia di rilevare che egli sembra aver avuto più «buona fortuna» che effettiva lucida comprensione delle difficoltà che doveva superare, concludendo che «il calcolo di Castillon è, da un punto di vista teorico, altrettanto errato di quello di Lambert, o, se l'erroneità ammette una gradazione, ancor di più. È assai probabile che egli procedesse in modo empirico, sì da evitare argomenti che dessero conclusioni non valide». Questo di Castillon si presenta come l'ultimo tentativo di elaborare un calcolo logico intensionale, secondo Couturat addirittura impossibile. Nell'opera citata, Lewis avanza l'ipotesi che probabilmente il «vuoto» verificatosi negli studi di logica nel periodo che intercorre fra Leibniz e Boole sia «dovuto alla predilezione per il punto di vista intensionale» e prosegue affermando che «non è casuale che gli inglesi abbiano avuto tanto successo in questo campo, una volta risvegliato il loro interesse per questa disciplina; essi pensavano abitualmente le relazioni logiche in estensione, e anche quando parlavano di "intensione"... essi non intendevano quelle relazioni di concetti intese con "intensione" nella logica tradizionale.»

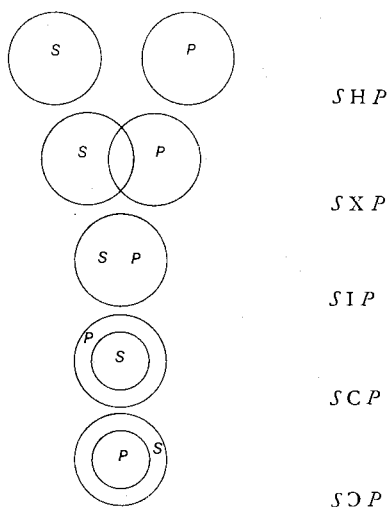
Ancora nel continente, si ha un interessante tentativo estensionale, ad opera del francese Joseph-Diez Gergonne (1771-1859) che pur occupandosi prevalentemente di geometria, dedicò tre saggi a problemi logici: *Essai de dialectique rationelle* (*Saggio di dialettica razionale*), *De l'analyse et de la synthèse dans le sciences mathématiques* (*Sull'analisi e la sintesi nelle scienze matematiche*) e *Essai sur la théorie des définitions* (*Saggio sulla teoria delle definizioni*), tutti pubblicati negli «Annales de mathématiques pures et appliquées» (di cui fu fondatore e direttore) rispettivamente i primi due nel tomo VII (1816-17) e il terzo nel tomo IX (1818-19). Torneremo fra poco sul primo dei tre saggi, che espone il calcolo estensionale. Nel secondo, Gergonne esprime la convinzione che tutte le definizioni sono nominali, e che la costituzione di una teoria si fonda sugli assiomi per quanto riguarda i teoremi e sui postulati per quanto riguarda i problemi; nel terzo saggio, di grande interesse anche se alquanto oscuro, Gergonne insiste sull'importanza del linguaggio simbolico per lo sviluppo delle scienze esatte e individua «... un tipo di pro-

posizioni che danno... la comprensione di uno dei termini di cui esse si compongono per mezzo del significato noto degli altri» e che «... potrebbero essere chiamate *definizioni implicite* in contrapposizione alle definizioni ordinarie che vengono dette *definizioni esplicite.* Si comprende così che... due proposizioni che contengono due parole nuove combinate con parole già note possono spesso determinare il senso delle prime». Ricordiamo infine che Gergonne, autore dallo stile duro e oscuro, sarà il primo a formulare in termini espliciti il principio di dualità per la geometria proiettiva, nel saggio *Considérations philosophiques sur les éléments de la science de l'étendue* (*Considerazioni filosofiche sugli elementi della scienza dell'esteso*) del 1826, grazie al quale si otteneva una *géométrie en parties doubles* che anticipa la sistemazione della proiettiva staccata da ogni riferimento metrico attuata in seguito da von Staudt.

Diversamente da Victor Poncelet (1788-1867) che nel suo *Traité des propriétés projectives des figures* (*Trattato delle proprietà proiettive delle figure*, 1822) aveva utilizzato il metodo basandosi su una corrispondenza punto-retta incentrata sul concetto di polare e reciproca, Gergonne ritiene che il principio non dipenda da particolari corrispondenze fra punto e retta ma discenda esclusivamente dal concetto stesso di *étendue.* In termini più espliciti Gergonne più di Poncelet ha la viva percezione del fatto che una volta considerata la geometria come teoria formale sul modello dell'algebra in cui le equazioni ammettono diverse soluzioni, hanno diverse interpretazioni che soddisfano le condizioni poste, le stesse proposizioni geometriche potevano valere tanto leggendo punto al posto di retta quanto viceversa, una volta che si considerassero anche i punti all'infinito. È così profonda in Gergonne la convinzione di poter estendere direttamente anche alla geometria l'approccio formale dell'algebra, che egli giunge a parlare esplicitamente per la prima volta di assiomi come definizione implicita dei modelli e quindi a porre, almeno in via embrionale, il problema della ricerca di modelli di assiomi dati in analogia alla ricerca di soluzioni di equazioni algebriche.

Sarà Julius Plücker (1801-68) nel 1828 a fornire una dimostrazione algebrica della dualità introducendo le coordinate omogenee, ma non molto più tardi Michel Chasles (1793-1880) e Christian von Staudt (1798-1867) mostreranno come la stessa geometria sintetica, una volta opportunamente liberata la concezione

degli assiomi da ogni riferimento univoco alla realtà, sia in grado di fornire la giustificazione non solo del principio di dualità, ma di altri principi di transfer che si basano su deformazioni e trasformazioni delle figure come il principio di continuità di Poncelet. Emergeva così a chiare lettere quell'implesso di motivazioni geometriche e logiche che portavano inevitabilmente a considerare il problema dei rapporti fra proposizioni e interpretazioni, della persistenza della verità di una proposizione passando da una interpretazione a un'altra. La logica assumeva quindi un ruolo particolare e nuovo e non stupisce che anche verso la logica tradizionale Gergonne si volga con uno spirito di carattere più matematico di quello usuale.

Nella *Dialectique rationelle*, ricollegandosi alla rappresentazione geometrica che dei sillogismi aveva dato Eulero, e riprendendo un'idea espressa da Condorcet che la riduzione aristotelica delle argomentazioni a schemi rigorosi e precisi poteva rappresentare il primo passo verso una non ancora raggiunta perfezione dell'arte del ragionare, Gergonne afferma infatti che tale perfezione avrebbe potuto essere ottenuta a condizione di riuscire a *meccanizzare* il processo argomentativo stesso. Escogita così una teoria nella quale ingloba i principi del sillogismo aristotelico sulla base di cinque relazioni fra le estensioni rispettive del predicato (P) e del soggetto (S) di una proposizione categorica:

Mediante le relazioni precedenti le proposizioni aristoteliche possono essere espresse come segue:

A) Ogni *S* è *P*	*S* I *P* *oppure* *S* C *P*
E) Nessun *S* è *P*	*S* H *P*
I) Qualche *S* è *P*	*S* X *P* *oppure* *S* I *P* *oppure* *S* C *P* *oppure* *S* Ɔ *P*
O) Qualche *S* non è *P*	*S* H *P* *oppure* *S* X *P*[3]

È evidente che le cinque relazioni definite da Gergonne sono esaustive e mutuamente esclusive ossia che date le estensioni (non vuote) di un predicato e di un soggetto comunque presi, certamente sussiste fra esse almeno una e al massimo una (e quindi una e una sola) delle relazioni suddette. Per quanto riguarda il calcolo pratico del sillogismo, considerato che le prime tre relazioni (H, X, I) sono simmetriche (ossia, ad esempio, che se vale *S* H *P* sussiste anche *P* H *S*) e che la penultima (C) è la conversa dell'ultima, è chiaro che si ottiene una grande semplificazione, perché è sufficiente assumere una qualunque «figura» aristotelica, ad esempio la prima

$$\frac{\begin{array}{c} M \ldots P \\ S \ldots M \end{array}}{S \ldots P}$$

e quindi inserire in tutti i modi possibili, negli spazi occupati dai tre punti, i segni delle cinque relazioni. Nel suo saggio Gergonne segue appunto questa procedura e riesce così a stabilire fra le 5^3 possibili sostituzioni che si possono effettuare, quali danno luogo a argomenti validi.

[3] Ecco come Gergonne giustifica la scelta dei simboli adottati per indicare le relazioni chiarite nel testo, ed esemplificate a sinistra, grazie ai cerchi di Eulero, come rapporti di inclusione. «I segni che caratterizzano queste relazioni sono stati scelti in quello che ci è sembrato il modo migliore per collegare il segno alla cosa significata, il che stimo di una certa importanza, per quanto a prima vista ciò possa apparire puerile. La lettera H, iniziale della parola *Hors* [al di fuori] designa il sistema di due idee completamente esterne l'una dall'altra, come appunto sono le due sbarre verticali di questa lettera. Queste due sbarre possono poi essere considerate incrociate per formare la lettera X con la quale si intende indicare il sistema di due idee che in qualche modo si intersecano. Infine le due sbarre possono coincidere sì da formare la lettera I che usiamo per rappresentare il sistema di due idee esattamente coincidenti l'una con l'altra; inoltre questa lettera

Per quanto ingegnoso e meritevole di essere ricordato a livello storico, il tentativo di Gergonne di rifondare la teoria del sillogismo si muove però su una linea che riprende idee settecentesche più che porsi come anticipatore della svolta booleana. Analogamente estraneo alla «linea booleana» ma tuttavia precursore di fecondissime idee maturate nel successivo sviluppo della logica contemporanea è il matematico, logico e filosofo cecoslovacco Bernard Bolzano (1791-1848); tanto nella sua *Wissenschaftslehre* (*Dottrina della scienza*) del 1837 quanto nei *Paradoxien des Unendlichen* (*Paradossi dell'infinito*), pubblicati postumi nel 1851, Bolzano introduce ed elabora tutta una serie di idee e concetti che risulteranno fondamentali nella sistemazione logica contemporanea. In particolare, nella prima delle opere citate giunge a una definizione di conseguenza logica che in parte coincide con quella oggi adottata dopo la sistematica e organica trattazione operata da Tarski nel 1936. Propriamente – come è stato posto in luce in tempi recenti – la definizione di Bolzano risulta limitata, rispetto a quella tarskiana, da due vincolanti. Per Bolzano ha senso domandarsi se *A* è conseguenza di *B* solo se *A* e *B* sono compatibili; inoltre Bolzano, pur ammettendo che il legame di conseguenza logica tra *A* e *B* coinvolge tutte le possibili interpretazioni delle componenti non logiche che compaiono nelle due proposizioni (nomi di individui, di proprietà, ecc.) non concepisce che anche i *domini* dei diversi universi in cui si interpreta possano variare e assume un universo unico. Quello che risulta estremamente interessante è la sicurezza con cui ad esempio in un lavoro posteriore alla *Wissenschaftslehre*, la *Grössenlehre*, rimasta incompiuta, Bolzano accuratamente separi dimensione sintattica da dimensione semantica, variabili da co-

è l'iniziale della parola *Identità*, denominazione opportuna per il tipo di relazione in questione. Si noti anche che le tre lettere H, X, I sono simmetriche proprio come le relazioni che esse debbono rappresentare, cosicché esse non cambiano il loro aspetto se sono rovesciate. Questo non avviene per la lettera C che una volta rovesciata diventa Ɔ; quindi abbiamo riservato questa lettera per indicare una relazione nella quale le due idee hanno un ruolo differente, una relazione che non è in generale reciproca. Questa lettera inoltre è l'iniziale comune di entrambe le parole *Contenente* e *Contenuto*, che esprimono adeguatamente la relativa situazione delle due idee». Si noti che quest'ultimo concetto di inclusione (che verrà poi indagato sistematicamente da Peirce nel 1870) compare qui esplicitamente per la prima volta nella logica moderna. Si osservi infine che «inclusione» significa evidentemente per Gergonne «inclusione propria».

stanti, simboli logici da extralogici. L'introduzione all'opera dal titolo *Von der mathematische Methode* (*Del metodo matematico*) mostra come Bolzano sapesse giustamente equilibrare aspetto sintattico e semantico sia nell'analisi delle dimostrazioni che nelle definizioni, giungendo ad introdurre distinzioni che secondo alcuni si ritrovano nella teoria della dimostrazione di Gentzen.

Se questo è vero per specifici risultati tecnici, non si può dire che il progetto generale di una teoria della struttura della scienza come concepita da Bolzano coincida con la metalogica di questo secolo. Bolzano è molto più attento a dimensioni del procedere scientifico (l'euristica, il concetto di evidenza e di spiegazione, la metodica della costruzione di trattati scientifici, ecc.) che danno alla sua indagine un orizzonte più vasto, anche se meno tecnicamente sviluppato di quello che noi oggi concepiamo. La capacità di ripensare dall'inizio la vera base della scienza la ritroviamo nei *Paradossi* ove egli sottopone ad acutissima analisi situazioni paradossali connesse con l'accettazione dell'infinito attuale, ponendosi (anche se la sua analisi presenta lacune ed errori talora ingenui) come diretto precursore delle riflessioni di Georg Cantor sulla teoria degli insiemi. Purtroppo Bolzano lavorava da isolato e non ebbe alcuna influenza (tanto in campo logico quanto in campo matematico, ove propugnava idee altrettanto innovatrici) sui contemporanei.

Fatta eccezione per l'opera generale di Bolzano sarebbe difficile, per delineare la situazione continentale, e in particolare tedesca, di questo periodo, per quanto riguarda la logica *formale*, trovare parole più esplicite e adeguate di quelle che John Venn ebbe a scrivere nell'introduzione alla *Symbolic logic* (*Logica simbolica*) del 1881 (e nelle quali è presente un riferimento al filosofo inglese Hamilton cui ci ricollegheremo in seguito). Dopo aver «confessato» il suo «spiacevole sospetto» circa l'influenza negativa del pensiero di Kant sullo sviluppo della logica formale, Venn prosegue affermando: «È comunque istruttivo confrontare il vigore e l'originalità con cui questa scienza [la logica] veniva trattata quando il grande filosofo non era ancora il "signor Immanuel Kant, professore di filosofia a Königsberg in Prussia", con la monotona inondazione di trattati di logica che invase la Germania per tanto tempo dopo di lui, e la cui ondata ci ha raggiunto attraverso le opere di Hamilton e di Mansel. Io ammiro profondamente la dottrina e

l'acutezza di molte di queste opere, elaborate in periodi in cui si poneva la massima cura nell'impedire l'intrusione della matematica, ma confesso che esse mi sembrano piuttosto ristrette se confrontate con ciò che venne prodotto quando lo spirito e la procedura della scienza sorella furono più liberamente accolti dai logici».

Non possiamo qui soffermarci – come si potrebbe fare – a «ridimensionare» in certo senso la responsabilità di Kant circa l'involuzione degli studi di logica formale determinatasi in Germania nella prima metà dell'Ottocento, involuzione – beninteso – per quanto riguarda l'aspetto appunto *formale* della logica. Tutto un altro discorso andrebbe fatto – ma non in questa sede in cui ci occupiamo esclusivamente della logica formale e matematica – per la logica nel senso generale. A nostro avviso nomi come quello di Fichte, Hegel, Herbart in Germania o di Mill, Hamilton, ecc. in Inghilterra hanno un peso rilevante nell'evoluzione della logica intesa come dottrina del sapere e dello sviluppo del pensiero. Forse verrà il giorno in cui la tradizione formale e (in mancanza di termini migliori) filosofica in senso stretto si potranno fecondamente incontrare. Quello che ci sembra importante ribadire è che si tratta di due approcci che hanno diversi meriti rispetto a diversi tipi di problematiche.

Per quanto riguarda la logica formale, resta comunque il fatto che malgrado Kant avesse riaffermato, in funzione antipsicologistica, la natura *formale* della logica, ne aveva fatto una scienza del «pensiero» e non del discorso, favorendo così la possibilità di una «ricaduta» in una sorta di psicologismo trascendentale, corroborata in effetti dalla contrapposizione, operata come sappiamo da Kant, di una logica trascendentale (quale teoria delle funzioni della conoscenza pura) alla logica formale pura. La successiva elaborazione idealistica della dottrina kantiana accentuerà a questo riguardo proprio l'aspetto psicologistico-trascendentale, risolvendo la logica kantiana esclusivamente nella sua parte trascendentale, che veniva poi interpretata come una «metafisica della mente» o dello «spirito».

È su queste basi che si fonda quella «inondazione» deprecata da Venn, e rendere conto della storia della logica continentale in questo periodo significherebbe o fare una vera e propria storia della filosofia generale oppure soltanto riportare una monotona e nu-

trita elencazione di titoli e nomi; basta scorrere ad esempio il *System der Logik und Geschichte der logischen Lehren* (*Sistema di logica e storia delle dottrine logiche*, 1857) di F. Überweg per rendersi conto del numero veramente enorme di scritti e manuali di logica di intonazione kantiana e idealistica che invase la Germania della prima metà dell'Ottocento. Naturalmente non mancarono oppositori all'interpretazione idealistica (e in particolare psicologistica) della logica; ad esempio Bolzano, che ripropone l'idea della logica formale come teoria delle *proposizioni in sé* (quindi in particolare indipendenti dal momento psicologico di essere pensate) e di un puro rapporto inferenziale da proposizioni a proposizioni; oppure Johann Friedrich Herbart (1766-1841) che definisce la logica come la scienza che considera in generale la chiarezza dei concetti e la loro combinazione in giudizi o inferenze ed assegna alla metafisica il problema del significato conoscitivo delle forme di pensiero sostenendo quindi che la logica non necessita, né è suscettibile, di una fondazione scientifica basata su considerazioni metafisiche e psicologiche. Ma queste e altre poche voci rimangono isolate e del tutto inascoltate in questo periodo di imperante filosofia idealistica; la «contaminazione» lamentata da Venn giunge in Inghilterra soprattutto attraverso l'opera di Hamilton e dei suoi allievi Veitch e Mansel, sotto forma essenzialmente di una interpretazione «psicologistica» della logica mutuata direttamente da questo carattere implicito nella dottrina kantiana.

Del resto la situazione in Inghilterra non è certamente più «consolante» in questo primo scorcio dell'Ottocento di quanto non lo sia nel Continente. Malgrado la logica fosse disciplina regolarmente inserita nei «piani di studio», ossia negli «statuti», delle università inglesi, essa veniva concepita come mera propedeutica e come tale veniva insegnata (anche se non in tutte). Ed è molto difficile, per non dire impossibile, indicare un autore o un'opera precisi che possano far presagire un nuovo orientamento nel senso di un rinnovato interesse per la logica formale. Ad esempio, due opere fondamentali come *A preliminary discourse on the study of natural philosophy* (*Discorso preliminare sullo studio della filosofia naturale*) di Johann William Herschel (1792-1871) pubblicato nel 1831, e la *History of the inductive sciences* (*Storia delle scienze induttive*) in due volumi editi nel 1840 da William Whewell (1794-1866) anche se riprendono una problematica genericamente di tipo logico, e rap-

presentano il rinascere di un interesse genuino per il metodo effettivo delle scienze, non vedono nella logica formale un momento centrale della ricerca. Un atteggiamento diverso è quello di John Stuart Mill (1806-73) che sviluppa una teoria logica che tenta di conciliare da una parte alcune classiche esigenze dell'empirismo dall'altra il riconoscimento del valore dell'argomentazione formale nella scienza e nella deduzione. In un certo senso Mill si pone quale mediatore tra «innovatori» come Herschel, legati alle scienze empiriche e che vedono nella logica formale un residuo del passato scolastico, e coloro i quali – come ad esempio Whewell – non rifiutano in nome dell'esperienza il riferimento al pensiero astratto, sino a giungere a forme di vero e proprio platonismo. È una lotta difficile che coinvolge il ruolo della matematica sul piano educativo e in cui gli appelli alla logica, come quelli ad esempio di Richard Whately (1787-1863) e dopo di lui del suo allievo, il futuro cardinale J.H. Newman (1805-90), autore di una *Grammar of Assent* (*Grammatica dell'assenso*), hanno il significato di una contrapposizione all'iconoclastia scientifica, che vuole sostituire la matematica alla logica come esempio del retto modo di procedere razionalmente non solo nelle scienze come la fisica, ma anche in campi quali l'economia politica.

In un senso, il grande ruolo che per gli algebristi della logica ebbe l'opera di Whately, è legato proprio al tentativo sviluppato negli *Elements of logic* (*Elementi di logica*, 1826) di mostrare come la validità logica avesse un fondamento esclusivamente linguistico e si basasse sul puro significato delle particelle logiche del discorso. È proprio in questa conciliazione fra formale e astratto che gli algebristi della logica vedono la possibilità di uno sviluppo matematico, che sia rispettoso dei contenuti specifici cui la logica può essere applicata. Il volume di Whately ha in effetti, oggi, solo valore storico e non offre contributi dal punto di vista dell'elaborazione formale. A dispetto comunque dei suoi limiti, esso agì da potente stimolo sugli autori inglesi. Tra il 1827 e il 1833 comparvero tra gli altri, direttamente o indirettamente influenzati da esso, una *Introduction to logic* (*Introduzione alla logica*, 1827) di Samuel Hinds, un *Outline of a new system of logic* (*Lineamenti di un nuovo sistema di logica*, 1827) di George Bentham (1800-84), un commento agli *Elements* di Whately, 1829, di Cornewall Lewis (1806-63) e un *Treatise of Logic* (*Trattato di logica*, 1833) di John Huyshe.

Sul piano più tecnico è per noi interessante notare il volume di George Bentham, dal quale citiamo direttamente il passo relativo a quella che verrà considerata, per la logica anteriore a Boole, la più rimarchevole e decisiva innovazione: l'introduzione della quantificazione del predicato. Nel capitolo VIII dell'*Outline* scrive dunque Bentham: «Nel caso in cui entrambi i termini di una proposizione siano entità collettive [concetti] può aver luogo identità o diversità: 1) Tra *ogni* individuo inteso dall'un termine e *ogni* individuo inteso dall'altro. Ad esempio: l'identità tra triangoli equilateri e triangoli equiangoli. 2) Tra *ogni* individuo inteso dall'un termine e *ognuno di una parte* soltanto degli individui intesi dall'altro. Ad esempio: l'identità tra uomini e animali. 3) Tra *ognuno di una parte* soltanto degli individui intesi dall'un termine e *ognuno di una parte* soltanto degli individui intesi dall'altro. Ad esempio l'identità fra quadrupedi e animali che nuotano... Le proposizioni semplici considerate rispetto alle relazioni precedenti possono, di conseguenza, essere o affermative o negative e ogni termine può essere universale o parziale. Di conseguenza queste proposizioni sono riducibili alle otto forme seguenti, nelle quali, allo scopo di astrarre da ogni idea non connessa alla sostanza di ogni specie, ho espresso i due termini con le lettere *X* e *Y*, la loro identità col segno matematico =, la diversità col segno $\|$, l'universalità con le parole *in toto*, la parzialità con le parole *ex parte*; o, per brevità, premettendo le lettere t e p come segni di universalità o parzialità. Queste forme sono:

1. *X in toto* = *Y ex parte* o t*X* = p*Y*
2. *X in toto* $\|$ *Y ex parte* o t*X* $\|$ p*Y*
3. *X in toto* = *Y in toto* o t*X* = t*Y*
4. *X in toto* $\|$ *Y in toto* o t*X* $\|$ t*Y*
5. *X ex parte* = *Y ex parte* o p*X* = p*Y*
6. *X ex parte* $\|$ *Y ex parte* o p*X* $\|$ p*Y*
7. *X ex parte* = *Y in toto* o p*X* = t*Y*
8. *X ex parte* $\|$ *Y in toto* o p*X* $\|$ t*Y*».

Bentham riduce successivamente a cinque le otto forme precedenti (il che dovrebbe risultare naturale al lettore, se tiene conto della analisi effettuata da Gergonne sulla base delle sue *cinque* relazioni esaustive e disgiunte) fra le quali segnala come particolarmente importante la t*X* = t*Y*.

Abbiamo riportato per esteso le parole di Bentham per rendere conto della gratuità della polemica sorta nel 1846 fra sir William Hamilton di Edimburgo (1788-1856) e Augustus De Morgan proprio sul tema della priorità della scoperta della quantificazione del predicato; polemica nella quale hanno notevole peso le idee di Hamilton circa i rapporti fra logica e matematica, sui quali si era pronunciato in un articolo del 1836 in risposta a un saggio di Whewell dedicato al ruolo della matematica nell'educazione umanistica. Circa la consistenza della polemica stessa, basta considerare che la quantificazione del predicato risaliva per lo meno a Leibniz, Ploucquet, Holland, Lambert; ma il momento grottesco sta nel fatto che Hamilton, che rivendica a sé la scoperta – da lui ritenuta fondamentale per una definitiva e soddisfacente sistemazione della logica – pur riconoscendo l'anticipazione dei continentali settecenteschi (che giudicava però casuale, e non valutata da questi in tutta la sua effettiva portata) ignora completamente la tabella di Bentham, che pure conosceva benissimo; e questa sua «dimenticanza» non viene certo giustificata dal fatto che effettivamente Hamilton, come vedremo, tenta di inquadrare questa tecnica in un piano organico generale. Comunque non mette conto di soffermarsi su questa disputa il cui solo aspetto positivo è forse quello – sopra ricordato – di aver stimolato Boole a pubblicare la *Mathematical analysis* nel '47.

William Hamilton, che può considerarsi l'ultimo rappresentante della scuola scozzese, fu professore di logica e metafisica all'università di Edimburgo, dove venne nominato nel 1836. Si hanno di lui delle *Lectures on Metaphysics and Logic* (*Lezioni di metafisica e logica*) in quattro volumi, pubblicati postumi i primi due nel 1860, gli ultimi due nel 1866 a cura di H.L. Mansel e J. Veitch. Nell'ambiente in cui doveva svilupparsi il germe della rinata logica formale, accanto a una sistemazione «algoritmica» della sillogistica tradizionale, che oggi conserva soltanto un mero valore storico, il nostro autore afferma una componente «filosofica» della logica, a carattere psicologistico, che egli matura sotto l'influenza di Kant. Nella convinzione che occorra «esprimere *esplicitamente* ciò che è pensato *implicitamente*», con l'impiego della quantificazione del predicato Hamilton ritiene di poter condurre a compimento l'ambizioso programma di fondare una «nuova analitica» che doveva servire a «completare e semplificare l'antica; a porre la chiave di

volta nell'arco aristotelico» e denuncia come uno degli «errori cardinali» dei logici che lo hanno preceduto proprio quello di non aver considerato che «nel pensiero il predicato ha sempre una quantità, esattamente come il soggetto, malgrado questa quantità venga frequentemente non enunciata in modo esplicito, come non necessaria nell'impiego ordinario del linguaggio; in quanto la nozione determinante o predicato viene sempre pensata come almeno adeguata a, o coestensiva con, il soggetto o nozione determinata, è raramente necessario esprimere questa quantità e il linguaggio tende a eliminare ciò che può essere omesso senza pericolo. Ma questa necessità interviene nel momento in cui, per conversione, il predicato diventa soggetto della proposizione; e ometterne la enunciazione formale significa degradare la logica da scienza della necessità del pensiero a un inutile accessorio dell'ambiguità delle parole».

Riportiamo ora, traendoli dalle *Elementary lessons in logic: deductive and inductive* (*Lezioni elementari di logica deduttiva e induttiva*) di William Stanley Jevons (di cui avremo occasione di occuparci in un prossimo paragrafo) la tabella delle otto proposizioni categoriche dovuta ad Hamilton, e alcuni esempi di sillogismi. Le lettere aggiunte alle designazioni medievali sono dovute a W. Thomson, un discepolo di W. Hamilton, che le presentò in *An outline of the necessary laws of thought* (*Lineamenti delle leggi necessarie del pensiero*) del 1849.

U	Tutti gli *S* sono tutti i *P*
I	Alcuni *S* sono alcuni *P*
A	Tutti gli *S* sono alcuni *P*
Y	Alcuni *S* sono tutti i *P*
E	Nessun *S* è nessun *P*
ω	Alcuni *S* non sono alcuni *P*
η	Nessun *S* è alcun *P*
O	Alcuni *S* sono nessun *P*

dove «tutti» va inteso in senso non distributivo («ogni») ma collettivo (giusto il punto di vista estensionale assunto da Hamilton) e analogamente «alcuni» non va inteso come «qualche» (che significa «almeno uno») bensì come «almeno alcuni e non tutti», ossia isola una parte propria dell'estensione del termine davanti al quale figura. È evidente nella sistemazione hamiltoniana (oltre alla coincidenza con la tabella di Bentham) il rapporto con quella di

Gergonne, costituita, come si ricorderà, sulla base di cinque relazioni. Anzi proprio da questo rapporto con le relazioni di Gergonne è possibile evidenziare difetti e limiti della tabella hamiltoniana, sui quali tuttavia non mette conto qui di insistere.

Ed ecco ora alcuni esempi di sillogismi.

I figura	tutti gli M sono tutti i P	M U P
	alcuni S non sono alcuni M	S ω M
	alcuni S non sono alcuni P	S ω P
II figura	tutti i P sono tutti gli M	P U M
	tutti gli S sono tutti gli M	S U M
	tutti gli S sono tutti i P	S U P
III figura	alcuni M sono alcuni P	M I P
	alcuni M non sono alcuni S	M O S
	alcuni S non sono alcuni P	S O P

Si può osservare che l'adozione della quantificazione del predicato semplifica la teoria classica della conversione e che d'altra parte il maggior numero di modi possibili che si ottengono nella deduzione sillogistica viene compensato dalla semplificazione delle «regole di inferenza», visto che il calcolo si fonda su un solo principio o «canone» che Hamilton esprime come segue: «La relazione peggiore che sussiste fra uno qualsiasi di due termini e un terzo termine comune con cui almeno uno dei due primi è in relazione positiva, sussiste anche fra quei due termini»[4] e che più semplicemente afferma che in un sillogismo una almeno delle premesse deve essere affermativa e che la conclusione «pende dal lato malvagio» ossia dal lato negativo e particolare.

Si noti che sulla base della contrapposizione fra il punto di vista della comprensione (o metafisico) e quello dell'estensione (o logi-

[4] Ove le relazioni, dalla «migliore» alla «peggiore», sono così ordinate: coinclusione toto-totale; coinclusione (e quindi coesclusione) incompleta; coesclusione toto-totale.

co) Hamilton esclude la quarta figura sillogistica ritenendola «non naturale», con argomentazioni assai affini a quelle che Kant aveva usato in uno scritto precritico del 1762, *Die falsche Spitzfindigkeit der vier syllogistichen Figuren* (*La falsa sottigliezza delle quattro figure sillogistiche*). Il richiamo a Kant non è casuale: ripresa ed elaborata la concezione kantiana della matematica come scienza della quantità fondata sulle forme pure dell'intuizione, Hamilton nel corso della polemica con De Morgan ritorna come dicevamo sulla questione dei rapporti fra logica e matematica riecheggiando temi sviluppati nell'articolo *On the study of mathematics* (*Sullo studio della matematica*) scritto nel 1836 in polemica con Whewell. Questi aveva sostenuto da una parte la superiorità della matematica sulla logica tradizionale e dall'altra aveva insistito sull'utilità della matematica quale strumento di formazione razionale nell'educazione umanistica.

Nel suo articolo Hamilton contrappone il «metodo filosofico» al «metodo matematico» e afferma che la matematica non ha il minimo valore educativo, dal momento che il simbolismo «blocca» il pensiero anziché educarlo; nel contempo stacca la logica dalla matematica inserendola nell'ambito della filosofia. A prescindere dalla sterilità di contrapposizioni del tipo di quella proposta da Hamilton, va osservato che questa sua presa di posizione avveniva in un contesto matematico tutto proiettato – come abbiamo visto – a disancorare i sistemi matematici da interpretazioni numeriche o genericamente «quantitative» in vista di una più ampia concezione del momento formale e precedeva di un solo decennio l'esplicita dichiarazione di Boole secondo la quale «il fatto che alle forme esistenti di analisi venga assegnata un'interpretazione quantitativa è il risultato delle circostanze che determinarono il sorgere di tali forme, e noi non dobbiamo farne una condizione universale dell'analisi».

Definita ora la filosofia come scienza della ricerca delle cause (e proprio per questo superiore alla matematica che «non sa nulla delle cause»), rifacendosi a temi kantiani Hamilton giunge a una sua determinazione come «scienza della mente»: «... poiché la filosofia non è puramente conoscenza, ma conoscenza delle cause e la mente stessa è la causa universale e principale concorrente in ogni atto di conoscenza, la filosofia deve fare della mente il primo e fondamentale oggetto della sua considerazione. Lo studio della

mente è così per eminenza lo studio filosofico.» La logica viene quindi ad essere un «frammento della scienza della mente» o, secondo la definizione preferita da Hamilton, «la scienza del pensiero in quanto pensiero». Donde risulta con evidenza il carattere «psicologistico» della logica hamiltoniana.

Sul significato del pensiero di William Hamilton per la logica, vale a dire sull'influenza che egli possa aver esercitato sugli sviluppi successivi, i pareri sono molto discordi. Ad esempio Lewis e Langford, riferendosi alla quantificazione del predicato ed estendendo poi in generale la loro valutazione, affermano: «L'idea [della quantificazione del predicato] è semplice, e in realtà di scarso rilievo per la logica esatta: di essa non è stato fatto alcun uso negli studi recenti.[5] Ma spesso l'importanza storica di un'idea dipende non tanto dai suoi meriti intrinseci quanto dallo stimolo esercitato sulle altre menti; e questo è il nostro caso. L'unico momento significativo della teoria della quantificazione del predicato per la logica esatta è che essa suggerisce un modo nel quale le proposizioni possono essere trattate come equazioni di termini; e la mera rappresentazione porta alla mente l'idea di un'analogia fra logica e matematica. Questo solo fatto congiunto con la fiduciosa assunzione di Hamilton che la logica stava per entrare in un nuovo periodo di sviluppo sembra essere stato un considerevole fattore nel rinnovamento degli studi logici in Inghilterra.» E Lewis aveva affermato, nel 1918, che «senza Hamilton avremmo potuto non avere Boole».

Barone viceversa – e di questo avviso siamo noi pure – concorda col precedente giudizio solo per quanto riguarda la scarsa importanza della quantificazione del predicato, ma non concede che l'analogia matematica suggerita nelle parole di Lewis e Langford abbia avuto un'effettiva influenza su Boole per quanto riguarda l'inserimento della ricerca logica in una dimensione schiettamente matematica. Egli sostiene invece che «a differenza dai temi hamiltoniani tecnico formali, la sua [di W. Hamilton] interpretazione filosofica della logica penetrò così a fondo nella cultura inglese del tempo da diventare quasi un tema obbligato

[5] Il volume di Clarence Irving Lewis e Cooper Harold Langford da cui è tratta la citazione è *Symbolic logic* (*Logica simbolica*) del 1932; ma quel «recente» può altrettanto bene essere riferito, da questo punto di vista, anche agli ultimi sviluppi della logica matematica. Il giudizio di Lewis riportato più avanti nel testo è tratto dal già citato *Survey* del 1918.

di essa. Così fu accolta anche da coloro che sul piano della logica come scienza procedettero in maniera originale e autonoma, ostacolando talvolta con le sue posizioni bloccate anche un libero sviluppo delle ricerche tecniche». E come conclusione di questa dialettica di sviluppo, osserva che «la logica formale moderna nasce in questa situazione travagliata; ma fu un travaglio fertile, perché da esso non soltanto venne un arricchimento della problematica scientifica, bensì ancor oggi si può trarne lo stimolo per ripensare criticamente il problema del significato dell'attività filosofica».

2. IL SORGERE DELL'ALGEBRA DELLA LOGICA: CONSIDERAZIONI INTRODUTTIVE

Abbiamo più volte accennato al fatto che la svolta decisiva che doveva impostare la logica su basi moderne avviene nel periodo a cavallo fra la prima e la seconda metà del XIX secolo: inizialmente e ufficialmente ad opera dell'irlandese George Boole, di cui appare, nel 1847, la fondamentale *The mathematical analysis of logic, being an essay towards a calculus of deductive reasoning* (*L'analisi matematica della logica: saggio di un calcolo del ragionamento deduttivo*). In effetti, dopo la pubblicazione di quest'opera, che già sappiamo essere strettamente collegata al processo di evoluzione subito nel primo scorcio del secolo dall'algebra in Inghilterra, venne sviluppandosi tutto un filone di ricerche sull'algebra della logica, che culminò verso la fine del XIX secolo e gli inizi del XX in una presentazione sistematica della materia, più avanzata e comprensiva rispetto al sistema originariamente esposto da Boole tanto nel lavoro testé ricordato quanto nel volume, del 1854, *An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities* (*Una ricerca sulle leggi del pensiero sulle quali sono fondate le teorie matematiche della logica e delle probabilità*), da lui considerato come il proprio capolavoro. A questa sistemazione rigorosa e organica si dedicò il tedesco Ernst Schröder con i tre volumi delle *Vorlesungen über die Algebra der Logik (exakte Logik)* (*Lezioni sull'algebra della logica – Logica esatta*) pubblicati tra il 1890 e il 1905, dopo che risultati significativi ed estensioni essenziali erano stati frutto delle fatiche di tutta una serie di ricercato-

ri europei e dell'americano Charles Sanders Peirce e della sua scuola.

In sede introduttiva possiamo limitarci semplicemente a enunciare tre elementi generali caratteristici della impostazione booleana della ricerca logica: 1) la *dimensione linguistica*, espressamente dichiarata da Boole, della ricerca stessa, alla cui identificazione ed esplicitazione concorre evidentemente la familiarità con la problematica della scuola degli algebristi di Cambridge: «Ciò che rende possibile la logica», dice Boole, «è l'esistenza nella nostra mente di nozioni generali, la nostra capacità di concepire una classe e di designare per mezzo di un nome comune gli individui che ne sono membri. La teoria della logica si rivela così intimamente connessa con la teoria del linguaggio.» Fin da ora va tuttavia aggiunto che il linguaggio è per Boole «uno strumento della logica, ma non... uno strumento indispensabile». 2) La *concezione psicologistica* della logica, che rappresenta il momento centrale, il punto di riferimento costante di tutta la speculazione logica di Boole. Da una parte infatti tale concezione «finalizza» il discorso logico di Boole, dall'altra si pone come garanzia epistemologica dello stesso, nel senso che è sempre un atto mentale la giustificazione ultima di un procedimento o di un'operazione logici. In certo senso ciò comporta che si può intendere l'indagine logica di Boole come una ricerca sui limiti e le possibilità della stessa attività simbolizzatrice dell'uomo. 3) L'espressa convinzione della *natura matematica* del processo logico inferenziale che, nel contesto di un'ampia discussione sul rapporto logica-matematica-filosofia in corso in Inghilterra ai suoi tempi, solo con Boole trova lo sbocco teoretico e operativo di una «svincolamento» della logica dal tradizionale ambiente filosofico di appartenenza per un suo inserimento *autonomo* fra le scienze matematiche.

Questo terzo punto può essere inteso in almeno due modi di diverso momento. Si può cioè concepirlo come l'asserzione di un qualche tipo di priorità della matematica rispetto alla logica, nel senso che la prima costituisce la fondazione della seconda, che è ad essa riducibile; oppure, in senso più circoscritto e operativo (si potrebbe dire: kantianamente) si può intendere con esso il fatto che la logica mutua dalla matematica tutti quegli elementi tecnici di rigore, di simbolizzazione e di procedure «risolutive» che le permettono di indagare in maniera più feconda e sistematica i

propri specifici problemi, senza che ciò comporti in alcun modo una «riduzione» della logica alla matematica. Indubbiamente da Boole tale riduzione è intesa in senso forte, vale a dire considerando appunto come matematica la «natura» stessa del processo logico: «Che la logica come scienza,» afferma Boole, «sia suscettibile di applicazioni molto vaste è ammesso; ma è ugualmente certo che le sue forme e processi ultimi sono matematici.» È ovvio che assumere quello che venne detto il «matematismo» di Boole intendendolo nel primo senso (forte) comporta quasi automaticamente l'accettazione anche del secondo aspetto; viceversa, intendendolo nel secondo senso non si è in alcun modo impegnati circa quella che può essere la natura del rapporto logica-matematica.

Con una sola eccezione significativa – costituita da Peirce – si può senz'altro affermare che gli autori della seconda metà dell'Ottocento che si muovono nell'ambito del filone booleano o «algebrico» della logica si applicano in generale alla elaborazione e allo sviluppo della parte puramente formale del calcolo di Boole, mentre trascurano di considerare il contesto più generale nel quale Boole inseriva le proprie riflessioni logiche. Così, per quanto riguarda il punto 1), se oltre a Peirce si escludono Stanley Jevons e Hugh McColl che dimostrano maggiore sensibilità al problema nella sua generalità, viene di fatto a cadere la questione del rapporto fra sfera logica e sfera linguistica; la ricerca addirittura febbrile di sempre nuove simbolizzazioni ha pregnanza essenzialmente pragmatica, di «buon uso» o maggiore perspicuità e comodità ed è in generale collegata più alla questione del matematismo di Boole che non alla problematica generale del rapporto di cui si diceva. Per quanto riguarda il punto 2), va considerato che lo sviluppo e la precisazione della sua concezione psicologistica sono per Boole la genuina motivazione del passaggio dall'opera del '47 a quella del '54; eppure questo contesto generale venne dai suoi continuatori o semplicemente trascurato, o etichettato come inutile e sovrabbondante; Schröder, che condivideva tale impostazione, la assunse in modo sostanzialmente acritico, sotto l'ulteriore influenza di tutto un discorso di marca psicologistica sulla logica che si svolgeva nella Germania della seconda metà dell'Ottocento. Non è casuale che fosse proprio Peirce, dato il suo acuto impegno filosofico, ad assumere una franca posizione sul problema, risolvendolo in modo diametral-

mente opposto; come dirà nella sua *Minute logic* (*Logica minuta*) «nessuno farà ingiustizia al presente trattato descrivendo la sua posizione come estremamente sfavorevole all'impiego della psicologia in logica».

Per quanto infine riguarda il «matematismo» di Boole, esso fu singolarmente oggetto di reiterati «attacchi» da parte dei suoi continuatori, i quali tuttavia in generale si diressero più contro la seconda accezione sopra elucidata, piuttosto che impostare il problema del rapporto in generale fra logica e matematica. Tipico a questo proposito è Stanley William Jevons (1835-82), un allievo di De Morgan, che sosteneva la propria opposizione in chiave quasi esclusivamente polemica, in nome di una sistemazione calcolistica della logica che fosse più aderente a un non meglio identificato «pensiero comune», nella quale quindi non fosse ammessa la non interpretabilità puntuale di ogni passaggio «matematico» del processo logico. Anche in questo caso, chi affronta il problema nella sua generalità è Peirce, che vi contribuisce con la critica originalità che gli è propria e che fa sì che egli possa considerarsi come l'anello di mediazione fra il filone algebrico della logica e quello «logicista» cui subito accenneremo.

Nello stesso torno di tempo infatti sul continente europeo lo sviluppo dell'indirizzo critico della matematica, la cui impostazione abbiamo visto risalire agli autori della prima metà del secolo come Cauchy, Abel, ecc., portava a risultati e nuovi temi di ricerca – quali l'aritmetizzazione dell'Analisi, la stessa nascita della teoria degli insiemi, la sistemazione della geometria o più in generale del metodo assiomatico – che favorivano grandemente, si potrebbe dire rendevano quasi «naturale», un ulteriore sforzo di sistemazione teorico-filosofica dei fondamenti della matematica. Questo compito venne assunto come programmatico dal tedesco Gottlob Frege (1848-1925) che ravvisò allo scopo l'esigenza di una preliminare sistemazione della logica, cui Frege appunto si accinse a partire dal 1879 e che sviluppò secondo una direzione indubbiamente diversa da quella booleana; solo superficialmente però tale sistemazione può essere considerata come ponentesi in radicale contrasto con quella di Boole (come a Frege fu rimproverato da numerosi critici): va detto invece che per molti versi essa è più sottile e profonda di quella del logico inglese e che nel complesso ciò è storicamente più che plausibile. Comunque sia – affronteremo la

questione nel prossimo capitolo – in questa nuova impostazione fregeana ritroviamo quell'ampiezza di orizzonti e di tematiche generali che descrivevamo come tipica di Boole rispetto alla quasi totalità dei suoi continuatori.

Si conviene appunto di chiamare «logicista» o «logicistico» questo filone di elaborazione logica, dal nome del programma fregeano di fondazione della matematica; passate inizialmente inosservate le opere di Frege, tale filone ebbe come «manifesto» completo e monumentale i tre volumi dei *Principia mathematica* (1910-13) di Bertrand Russell (1872-1970) e di Alfred North Whitehead (1861-1947). Come abbiamo accennato, nelle sue implicazioni generali esso era stato in qualche senso precorso da alcuni autori di scuola algebrica, quali ad esempio Hugh McColl e soprattutto Peirce; quest'ultimo in particolare, che abbiamo già indicato come vero e proprio elemento di passaggio fra i due filoni, mostra che la rigida contrapposizione fra di essi è frutto in generale di pregiudizi; e che l'unico modo per evitare di impastoiarsi in classificazioni del tutto accidentali ed estrinseche consiste nell'affrontare una data problematica con la mente libera da chiusure o «etichette» precostituite. Malgrado quindi il filone algebrico o quello logicista vengano di norma contrapposti l'un l'altro (e tali apparvero, in effetti, a molti dei protagonisti della nostra storia), indipendentemente dagli indirizzi o dalle convinzioni personali dei vari autori è invece possibile riconoscere il sussistere fra di essi di un rapporto di reciproco completamento che è pienamente emerso nello sviluppo della logica formale nel nostro secolo (come vedremo in un prossimo paragrafo).

3. L'OPERA DI AUGUSTUS DE MORGAN

La posizione di Augustus De Morgan[6] nella storia della logica, e in particolare nei riguardi del suo più celebre contemporaneo Boole, non è qualificabile in modo univoco. Il comune giudizio

[6] Nacque nel 1806 a Madura, in India, come quinto figlio del colonnello John De Morgan, ufficiale al servizio della East India Company. Condotto in Inghilterra durante l'infanzia, perdette il padre a 10 anni e fu educato a una vasta e profonda conoscenza dei classici, che tuttavia non valse certo a mettere in luce la naturale

storiografico in proposito vede in De Morgan esclusivamente un sistematore della logica classica più che un vero e proprio innovatore (quale è appunto Boole). A nostro parere, pur essendo nella sua essenza corretto, tale giudizio è alquanto sbrigativo e parziale; nella vastissima produzione di De Morgan è infatti possibile individuare motivi che, oltre a confermarlo, da un lato mettono in evidenza la sua funzione di «precursore» nei riguardi di Boole, dall'altro mostrano invece l'influenza su di lui esercitata dal più illustre contemporaneo e, per certi versi, ne fanno paradossalmente addirittura un «prosecutore» del disegno generale espresso da Boole.

Da una parte infatti, e soprattutto per i suoi contributi all'algebra nel contesto della scuola di Cambridge, nonché per alcuni concetti introdotti nelle successive sistemazioni che egli elaborò per la teoria sillogistica, si può con Barone vedere in De Morgan «l'ultimo anello di quella catena che congiunge il Woodhouse con Boole» e sottolineare il momento di De Morgan «precursore», in particolare, ovviamente, nei suoi scritti precedenti il 1847. D'altra parte è pur vero che malgrado il livello personale cui giunse la controversia con W. Hamilton[7] e pur con tutta la preoccupazione

inclinazione e il non comune talento matematico del giovane. A 17 anni entrò nel Trinity College, dove abbandonò lo studio dei classici per letture matematiche e filosofiche; qui conobbe, tra gli altri, Airy, Peacock e Whewell. Una sua eventuale carriera accademica a Cambridge venne bloccata per motivi religiosi; ma nel 1828 venne chiamato alla libera University of London (l'attuale University College) che veniva appunto fondata in quell'anno, come professore di matematica; restò in questo posto per tutta la vita salvo alcuni periodi (ad esempio nel 1831 per cinque anni e quindi nel 1853) durante i quali egli rinunciò alla cattedra essenzialmente ancora per motivi religiosi. Ottimo insegnante, esperto consulente di compagnie di assicurazione, spirito indipendente e individualista, rifiutò l'associazione alla Royal Society e alla British Association. Fu scrittore estremamente fecondo: oltre a numerosi volumi, scrisse oltre 700 articoli per varie riviste (noi ci limiteremo a citare nel testo, della sua produzione, quanto più direttamente interessa la nostra esposizione). Morì nel 1871.

[7] Della disputa fra W. Hamilton e A. De Morgan si è già accennato nel paragrafo precedente. A parte il giudizio di Barone secondo il quale questa polemica «servì a sgretolare una tradizione secolare di pensiero e a preparare e favorire il rinnovamento booleano», si era detto in quella sede che in effetti la discussione, che tra l'altro verteva su un argomento (la priorità della scoperta della quantificazione del predicato) dimostratosi poi di importanza del tutto trascurabile per l'effettivo sviluppo della logica, era scaduta a un livello di scontro puramente personale fra i due contendenti e aveva di fatto perduto ogni contenuto significativo. Va tuttavia notato che essa agì anche come grosso stimolo alla produzione

del nostro autore di distinguersi dal suo avversario, in questo campo De Morgan deve essere giustamente considerato, proprio al pari di Hamilton, un sistematore della vecchia logica piuttosto che un fondatore della nuova; anche in questo caso tuttavia egli introdusse concetti ripresi poi da Boole e ne subì peraltro l'influenza riscontrabile direttamente in alcuni dei suoi lavori.

Infine, in vista della sua considerazione di una logica generale delle relazioni (o meglio della «possibilità» di tale logica implicita nel suo discorso) De Morgan, almeno a livello di spunti e anticipazioni, va considerato in posizione più avanzata rispetto a Boole, come suo «prosecutore», per dirla in una sola parola. Mentre quest'ultimo infatti, nel processo inferenziale inteso estensionalmente, si limiterà come vedremo a considerare il classico rapporto di inclusione fra classi, resterà cioè ancorato alla classica forma soggetto-predicato della proposizione, De Morgan propone invece di considerare le *proprietà* delle relazioni stesse, da una parte gettando così le basi di una logica delle relazioni, dall'altra, per quanto in particolare riguarda il momento inferenziale, proponendosi di considerare il rapporto sillogistico quando al posto della copula «è» (o «sono») venisse appunto sostituita una relazione

logica di De Morgan: si può dire che dal 1846 in poi non vi sia sua opera nella quale in modo diretto o indiretto non si riscontri un preciso riferimento alla polemica in questione. Comunque, per dare al lettore un'idea del «tono» della disputa, riportiamo il brano di una lettera che Hamilton, indirizzava al suo avversario il 13-3-1847 con l'evidente «odiosa accusa di plagio» poi lamentata da De Morgan: «... Sembra che Lei rivendichi a se stesso la scoperta indipendente della teoria fondamentale del sillogismo che io Le avevo comunicato privatamente... Non posso ammettere questa pretesa, anche se intesa a rivendicare una originalità di seconda mano. Per me è manifesto che per quanto riguarda il principio della teoria Lei sia completamente indebitato con me; e non posso fare a meno di pensare che se Lei presenta questa teoria come frutto di una speculazione del Suo pensiero (anche se riconosce a me la priorità sulla questione), ebbene Lei si rende colpevole – mi perdoni la franchezza – di un abuso di fiducia nei miei riguardi e di condotta sleale nei riguardi del pubblico». De Morgan risponde immediatamente per rigettare «l'odiosa accusa» con una lettera nella quale fra l'altro si legge: «Mi pregio di informarLa che attenderò fino al 10 del mese prossimo per una delle due cose: o una Sua ritrattazione o l'annuncio preciso del tempo e del luogo in cui Ella vorrà sostenere in pubblico la verità dell'accusa che – mi scusi l'espressione – Ella ha osato avanzare contro di me. Se per la data sopra specificata non avrò ricevuto una presa di posizione da parte sua in uno dei due sensi sopra citati, procederò immediatamente a stendere una dichiarazione, della pubblicazione della quale, non occorre dirlo, Ella sarà debitamente avvertito.» Crediamo possa bastare per esemplificare il livello su cui si manteneva tale controversia «scientifica».

qualunque soddisfacente certe proprietà generali. In quest'ultimo punto a nostro parere risalta in modo particolare l'originalità del pensiero di De Morgan e nel contempo il suo condizionamento di fondo, la sua cronica incapacità cioè a staccarsi in modo definitivo e fecondo dall'archetipo per eccellenza del processo inferenziale, il sillogismo aristotelico.

L'opera «algebrica» di De Morgan si esplica sostanzialmente con gli *Elements of algebra* (*Elementi d'algebra*, 1837), con quattro memorie pubblicata fra il 1839 e il 1844: *On the foundations of algebra* (*Sui fondamenti dell'algebra*) I, II, III e IV (quest'ultima avente pure l'ulteriore titolo *On triple algebra* [*Sull'algebra triplice*]), con un trattato *Differential and integral calculus* (*Calcolo differenziale e integrale*) del 1842, cui può aggiungersi il volume del 1849 dal titolo *Trigonometry and pure algebra* (*Trigonometria e algebra pura*). È soprattutto la serie di articoli sui fondamenti dell'algebra che illustra lo sfondo culturale e filosofico su cui si pongono non solo i contributi di De Morgan alla logica, ma l'intero sviluppo dell'algebra della logica britannica. In questione è la natura della matematica nel suo complesso: il contrasto tra aspetto descrittivo e aspetto formale ed operativo, il suo valore culturale e pedagogico, i rapporti fra geometria sintetica (non si dimentichi che nell'Inghilterra vittoriana a livello educativo matematica voleva dire essenzialmente *Elementi* di Euclide) e geometria analitica. Come posto in luce da studi recenti (ad esempio da Joan Richards in *Mathematical Visions* [*Visioni matematiche*] del 1988) è su questo terreno che si combatté la lotta tra coloro che volevano conservare all'insegnamento e alla ricerca matematica vitalità e valore culturale e quelli che ne vedevano solo il puro aspetto applicativo. In questi dibattiti fu proprio l'arcivescovo Whately – che pure tanto De Morgan che Boole salutarono come un restauratore degli studi logici – a sostenere il valore propedeutico e formativo della logica *contro* la matematica. Dall'altra parte, lo schieramento comprendeva personaggi come Whewell, Herschel, Boole, ecc. e tutti coloro che – su basi filosofiche diverse – vedevano la necessità di una differente integrazione della matematica nella vita intellettuale e della matematica inglese in quella europea. Su questo sfondo la scoperta che la logica si poteva concepire come scienza formale matematica acquista una rilevanza che è difficile sopravvalutare.

In linea con le concezioni della scuola «filologica», in queste

opere De Morgan riconosce l'equivocità del linguaggio, la necessità di separare l'aspetto logico-formale da quello semantico-contenutistico o, in termini già noti, l'algebra simbolica dalle sue applicazioni. Evidente è anche, nella quarta delle memorie su ricordate, l'influenza del matematico W.R. Hamilton e l'interesse di De Morgan per la costituzione tecnica di nuovi sistemi algebrici. Il momento originale rispetto ai canoni della scuola di Cambridge, e per noi particolarmente interessante, è l'esigenza da lui esplicitamente dichiarata di un'analisi logica dell'algebra, a proposito della quale egli distingue fra 1) *algebra logica* che è «la scienza che investiga il modo di dare significato ai simboli primitivi e di interpretare tutti i conseguenti risultati simbolici» e 2) *algebra tecnica* che sarebbe invece «l'arte di usare i simboli in base a regole che sono... prescritte come le definizioni dei simboli»; sulla base di questa distinzione sembra possibile concludere che mentre De Morgan sottolinea il carattere ipotetico-deduttivo e formale *all'interno* del sistema algebrico, ne affida il carattere «scientifico» e conoscitivo giusta la 1), all'aspetto esterno, exstrasistematico.

L'algebra della logica di De Morgan viene così a trovarsi posta – come sottolinea Barone – «di fronte al problema più generale del simbolismo e dell'attività della mente umana da cui esso dipende», per cui sfocia in effetti in una sorta di «operativismo» che gli fa peraltro presagire con assoluta chiarezza la comparsa di «altre algebre o almeno... l'estensione di quella attuale»: tra le quali, appunto, la booleana algebra della logica o i quaternioni di Hamilton.

La maggior parte della produzione logica di De Morgan è dedicata a successivi messe a punto e ampliamenti della sillogistica tradizionale. Ci limiteremo qui a ricordare sei fondamentali memorie dal titolo comune *On the syllogism* (*Sul sillogismo*) I, II, III, IV, V, VI apparse fra il 1846 e il 1868, ognuna con un suo proprio sommario che specificava l'aspetto o il tema particolare in essa trattato, il volume *Formal logic* (*Logica formale*) del 1847, che De Morgan stesso dichiarò essere apparso nelle librerie nel medesimo giorno della *Mathematical logic* di Boole;[8] la voce *Logic* (*Logica*) del 1860, scritta per il quinto volume della English Cyclopaedia. Dello stes-

[8] Così ricordò De Morgan quando, nel 1867, comunicò alle *Transactions* di Cambridge l'opera postuma di Boole *On propositions numerically definite* (*Sulle proposizioni numericamente definite*) che venne poi pubblicata nel volume XI (1868).

so anno infine va ricordato il *Syllabus of a proposed system of logic* (*Sommario di un sistema proposto di logica*) che egli stesso presenta come «summa» di tutta una serie di lavori precedenti, e al quale ci siamo in generale riferiti nel seguito per il simbolismo adottato. La terza delle memorie sul sillogismo sopra ricordate e il saggio *On the symbols of logic, the theory of syllogism and in particular of the copula, and the application of the theory of probabilities to some questions of evidence* (*Sui simboli della logica, la teoria del sillogismo e in particolare della copula, e l'applicazione della teoria delle probabilità ad alcune questioni di evidenza*) del 1851, sono due fra i lavori che più mostrano l'influenza su De Morgan di una problematica più specificamente booleana.

Nella sua sistemazione della sillogistica De Morgan introduce alcune idee che indubbiamente agirono su Boole e furono da questi riprese, come ad esempio l'impiego dei termini positivi e negativi («contrari» nella terminologia demorganiana), il concetto di universo del discorso e l'esplicazione delle cosiddette leggi di De Morgan. Si è già detto peraltro che De Morgan limita sostanzialmente l'impiego di queste innovazioni a una sistemazione definitiva del sillogismo; e tuttavia si muove, si può dire, in senso moderno, con una logica riguardata come scienza formale che direttamente nulla ha a che fare con la metafisica e la psicologia,[9] e le cui forme sono forme di un pensiero possibile piuttosto che attuale.

Dicevamo che le novità importanti che staccano De Morgan dalla tradizione classica sono l'introduzione dei termini negativi, che vengono indicati con la stessa lettera del corrispondente termine positivo, ma minuscola (cosicché, ad esempio, se X indica «razionale», x indicherà «non-razionale») e l'introduzione dell'universo del discorso, concetto questo strettamente legato al punto precedente nel senso che un termine X e il suo contrario x

[9] Ecco infatti come inizia il citato *Syllabus*: «1. La logica analizza le *forme*, o le *leggi di attività* del pensiero. 2. La logica è *formale*, non *materiale*: essa considera la legge di attività indipendentemente dalla materia su cui agisce. Essa non è *psicologica* né *metafisica*: essa non considera né la mente in se stessa, né la natura delle cose in se stesse; ma la mente in relazione alle cose e le cose in relazione alla mente. Cionondimeno, essa è *psicologica* nella misura in cui si occupa dei risultati della costituzione della mente; ed è *metafisica* nella misura in cui si occupa del retto uso di nozioni circa la natura e la dipendenza delle cose che, siano esse vere o false, in quanto rappresentazioni dell'esistenza reale intervengono nei modi comuni di pensare di tutti gli uomini.»

esauriscono questo universo, qualunque esso sia; De Morgan ammette inoltre che l'universo in questione non sia vuoto, abbia cioè almeno un oggetto (o, altrimenti detto, in senso leggermente diverso: una almeno di due categorie contrarie non è vuota). La portata di queste innovazioni è per lo meno duplice: viene eliminato il classico problema relativo alle «non entità», alle classi vuote, e superata la questione circa l'importo esistenziale delle proposizioni universali; scompare esteriormente la distinzione fra proposizioni affermative e negative (sicché De Morgan dovrà introdurre un'apposita regola per distinguerle).[10] L'introduzione dei termini negativi permette immediatamente di scrivere otto proposizioni categoriche invece delle classiche quattro (A, E, I, O):

Universali:
1. Tutti gli X sono Y
2. Tutti gli x sono y
3. Tutti gli X sono y
4. Tutti gli x sono Y

Particolari:
5. Alcuni X sono Y
6. Alcuni x sono y
7. Alcuni X sono y
8. Alcuni x sono Y

Per quanto ora riguarda il simbolismo impiegato da De Morgan, viste le successive modificazioni cui egli lo sottopose, noi ci riferiremo, come già detto, alla versione del *Syllabus* che può sostanzialmente considerarsi come definitiva. Malgrado le numerose variazioni, comune a ogni versione del simbolismo demorganiano è il fatto che viene indicata la distribuzione[11] di ogni termine e la

[10] Come vedremo fra un momento, simbolicamente una proposizione negativa viene indicata con un punto opportunamente disposto. Ecco come De Morgan, nel *Syllabus*, chiarisce la questione, definendo «*affermativa* una proposizione vera di X e X, falsa di X e non-X o x; *negativa*, una proposizione vera di X e x, falsa di X e X». Nota quindi che una volta introdotti i termini negativi «è impossibile definire l'opposizione di qualità mediante l'affermazione della negazione: poiché ogni affermazione è una negazione e ogni negazione è un'affermazione. La negazione di "Nessun X è Y" è l'affermazione "Tutti gli X sono y". La necessaria distinzione fra *affermativa* e *negativa* è tuttavia data sopra nel testo...» vale a dire consiste nella definizione sopra riportata.

[11] De Morgan non parla di «distribuzione», bensì usa locuzioni quali «termine universalmente predicato di...» o «particolarmente predicato di...» (ove noi di-

qualità, affermativa o negativa, di ogni proposizione. Per indicare che un termine X è distribuito egli scrive una parentesi (*spicula*) prima o dopo il termine stesso con la concavità rivolta verso di esso: $X)$ oppure $(X$; per indicare invece un termine non distribuito, la parentesi volge la convessità verso il termine: $X($ oppure $)X$. Così ad esempio $X))Y$ indica una proposizione il cui soggetto, X, è distribuito e il cui predicato, Y, non lo è, vale a dire «ogni X è Y»; mentre ad esempio $X()Y$ indica una proposizione con soggetto e predicato non distribuiti, ossia «alcuni X sono Y». Per indicare una proposizione negativa, fra le due parentesi che figurano nella proposizione stessa viene posto un punto «.». Ad esempio, «alcuni X non sono Y» verrà simbolizzata con $X(.(Y$ (si badi bene che *non* si tratta della negazione della $X(\ (Y$ corrispondente). Due punti equivalgono a nessun punto e indicano una proposizione affermativa. Nel simbolismo descritto si ha, per le otto proposizioni sopra indicate,

1. $X))Y$	5. $X()Y$
2. $x))y$	6. $x()y$
3. $X))y$	7. $X()y$
4. $x))Y$	8. $x()Y$

Si possono eseguire trasformazioni equivalenti su una proposizione secondo la seguente regola: cambiare la distribuzione di un termine – vale a dire rovesciare la parentesi ad esso relativa – cambiare un termine nel corrispondente negativo e cambiare la qualità della proposizione; ci si convince facilmente della validità di questa regola sulla base della legge della doppia negazione (*duplex negatio affirmat*) e di proprietà elementari della distribuzione: si ha allora subito che ognuna delle otto proposizioni precedenti ammette quattro forme equivalenti come mostrato dalla tabella che segue:

remmo «distribuito» o «non distribuito» rispettivamente), oppure parla semplicemente di «quantità» di un termine. Sorvoliamo qui su alcune precisazioni che a rigore andrebbero fatte circa il concetto di distribuzione; nel senso qui impiegato, intuitivamente un termine è distribuito quando indica la classe cui *tutti* i suoi elementi appartengono.

	(*a*)		(*b*)		(*c*)		(*d*)
1.	$X))Y$	=	$X).(y$	=	$x((y$	=	$x(.)Y$
2.	$x))y$	=	$x).(Y$	=	$X((Y$	=	$X(.)y$
3.	$X))y$	=	$X).(Y$	=	$x((Y$	=	$x(.)y$
4.	$x))Y$	=	$x).(y$	=	$X((y$	=	$X(.)Y$
5.	$X()Y$	=	$X(.(y$	=	$x)(y$	=	$x).)Y$
6.	$x()y$	=	$x(.(Y$	=	$X)(Y$	=	$X).)y$
7.	$X()y$	=	$X(.(Y$	=	$x)(Y$	=	$x).)y$
8.	$x()Y$	=	$x(.(y$	=	$X)(y$	=	$X).)Y$

Si noti che ogni riga della tabella, costituita da proposizioni fra loro equivalenti, contiene una proposizione con entrambi i termini positivi. Assumendo tali proposizioni come rappresentanti delle altre della stessa riga, abbiamo le otto proposizioni fondamentali

1. $X))Y$ Tutti gli *X* sono *Y*
2. $X((Y$ Alcuni *X* sono tutti gli *Y* [oppure, Tutti gli *Y* sono *X*]
3. $X).(Y$ Nessun *X* è *Y*
4. $X(.)Y$ Ogni cosa [dell'universo del discorso] è *X* o *Y* (o entrambi)
5. $X()Y$ Alcuni *X* sono *Y*
6. $X)(Y$ Alcune cose [dell'universo del discorso] non sono né *X* né *Y*
7. $X(.(Y$ Alcuni *X* non sono *Y*
8. $X).)Y$ Tutti gli *X* non sono alcuni *Y* [oppure, Alcuni *X* sono *Y*].

È superfluo soffermarci su alcune difficoltà che sorgono in questa sistemazione della sillogistica (il lettore, ad esempio, avrà senz'altro notato la «singolare» interpretazione della 4, che come si rileva dalla tabella delle equivalenze è una universale negativa, e della 6, che è invece una particolare affermativa, interpretazione dovuta alla necessità di conservare alcuni rapporti inferenziali). Il «quadrato» delle opposizioni diventa nel caso presente

1 contraddice 7   2 contraddice 8

3 contraddice 5   4 contraddice 6

vale a dire sono contraddittorie due proposizioni aventi gli stessi

termini se una è affermativa, l'altra negativa e la distribuzione dei termini nelle due proposizioni è opposta. A parte dicevamo alcune osservazioni particolari, quello che va messo in evidenza è che con questo simbolismo De Morgan ottiene una chiara sistemazione del sillogismo tanto con proposizioni singole (sillogismo unitario) quanto con sistemi di proposizioni (sillogismo cumulativo). Gli schemi fondamentali di deduzione sono i seguenti quattro:

$$1 \;\; \frac{))\quad))}{))} \qquad\qquad 3 \;\; \frac{((\quad()}{()}$$

$$2 \;\; \frac{()\quad))}{()} \qquad\qquad 4 \;\; \frac{((\quad((}{()}$$

Prendendo i tre termini *X*, *Y*, *Z* e i loro negativi *x*, *y*, *z* nelle otto differenti combinazioni *XYZ*, *xYZ*, *xyZ*, *xyz*, *xYz*, *XYz*, *Yyz*, *XyZ*, i quattro schemi precedenti danno, dei 64 sillogismi possibili (che si otterrebbero considerando anche gli schemi (), ((,)) (), ecc.) i 32 validi, e questi vengono così distinti: 8 universali (premesse e conclusione universali), 8 minori-particolari (premessa minore [la prima] particolare e conclusione particolare), 8 maggiori-particolari (premessa maggiore [la seconda] particolare e conclusione particolare), 8 particolari rafforzati (premesse universali e conclusione particolare). La regola con la quale vengono esclusi i 32 sillogismi non validi è la seguente: entrambe le premesse debbono essere universali o, se lo è solo una, il termine medio deve essere distribuito una sola volta (vale a dire il termine medio deve avere differenti quantità nelle due premesse). La regola operativa è invece la seguente: cancellare il termine medio e le parentesi ad esso relative.

Si avrà così, ad esempio

$$\frac{X((Y \quad Y()Z}{X()Z}$$

o ancora

$$\frac{X).(Y \quad Y(.)Z}{X)..)Z \text{ ossia } X))Z}.$$

I vantaggi di questo metodo rispetto ai sistemi tradizionali sono

molteplici: essendo indicata la distribuzione diventa superflua la conversione *per accidens* e la distinzione fra figure; si ottiene inoltre l'eliminazione di alcune forme ridondanti. De Morgan riesce anche a eliminare alcune difficoltà riconosciute nell'elaborazione hamiltoniana. Ma non conviene insistere su questo né è opportuno soffermarci a evidenziare limitazioni e difetti di questo sistema, che peraltro, si badi bene, si muove in pieno «clima» aristotelico. De Morgan infatti va al di là della sillogistica tradizionale solo quando tratta tipi particolari di sillogismo, come ad esempio i sillogismi composti, o i sillogismi numericamente definiti o infine i sillogismi dell'asserzione indecisa e della quantità trasposta.[12]

Ma il superamento essenziale ottenuto da De Morgan in questo senso dipende dall'aver riconosciuto che la copula del sillogismo tradizionale esplica la sua funzione non tanto per sue «più o meno misteriose» proprietà intrinseche, ma essenzialmente in virtù del fatto che, in quanto relazione, gode delle proprietà generali di simmetria e transitività. De Morgan ne conclude che ogni relazione che goda delle stesse proprietà potrebbe servire allo scopo e la realizzazione di questo sillogismo «generalizzato» è lo spunto iniziale alla sua impostazione di una teoria generale delle relazioni. È proprio in questo caso – lo abbiamo già accennato – che si può cogliere con tutta chiarezza la funzione stimolante e nel contempo limitativa che il costante riferimento alla sillogistica aristotelica ha esercitato su De Morgan. A partire da essa infatti egli riesce a impostare *in nuce* una teoria delle relazioni che tuttavia, malgrado egli ne intravveda una autonoma possibilità di sviluppo, ritorna ancora ad essere bloccata in funzione dello scopo fondamentale e privilegiato di costituire una *sillogistica* generalizzata.

[12] Coi sillogismi composti, De Morgan riduce a un tipo di calcolo sostanzialmente analogo al precedente relazioni composte fra proposizioni, nel senso che sono definite in base al sussistere fra di esse di due delle relazioni precedentemente considerate; così ad esempio introduce in particolare la relazione simbolizzata $X \| Y$ per indicare $X))\ Y$ e $Y\ ((X$ (ossia «tutti gli *X* sono *Y* e tutti gli *Y* sono *X*»); nei sillogismi della quantità trasposta secondo le sue stesse parole «l'intera quantità di un termine della conclusione, o del suo contrario, è applicata in una premessa all'altro termine della conclusione o al suo contrario (alcuni *X* non sono *Y*; *per ogni X* esiste un *Y* che è *X*; quindi alcuni *Z* non sono *Y*)». Nei sillogismi numericamente definiti o dell'asserzione numerica, una determinazione numerica o di probabilità si trasferisce dalle premesse (o da una delle premesse) alla conclusione.

Malgrado tuttavia questo orizzonte bloccante, determinato dall'aspirazione di ottenere una visione più astratta del ragionamento sillogistico[13] al fine di stabilirne un modello generale da cui potessero essere ottenute le forme aristoteliche e molte altre, De Morgan riesce a stabilire una serie di leggi generali valide nell'ambito di un'autonoma teoria delle relazioni, al punto che Peirce, cui si deve, come vedremo, una prima sistemazione di questo tipo verso la fine del secolo, riconoscerà in lui il «padre della logica dei relativi». Anche il simbolismo escogitato da De Morgan per questa sua generalizzazione soffre di numerose ambiguità, anche in dipendenza proprio dalla finalità cui tutta la sua ricerca è condizionata. La più grave di tali ambiguità resta comunque quella di impiegare una stessa lettera per indicare la relazione e gli elementi della relazione stessa. Nei limiti del possibile, tuttavia, nella breve esposizione che segue noi ci atterremo al simbolismo demorganiano originale, da lui presentato nella quarta delle sei note sul sillogismo sopra ricordate, che appunto porta il titolo *On the syllogism,* IV *and on the logic of relations* (*Sul sillogismo,* IV *e sulla logica delle relazioni*), del 1860.

In questo contesto, X, Y, Z rappresentano termini singolari o nomi di classi; L, M, N, relazioni. $X..LY$ significa che X è un L di Y o, più precisamente, che X è qualcuno degli oggetti del pensiero che stanno con Y nella relazione L, ossia che è uno degli L di Y (si pensi ad esempio che L significhi la relazione X è amante di Y dove X e Y vanno intese come classi). $X.LY$ nega la relazione precedente ossia afferma che X non sta nella relazione L con Y (X non è amante di Y). $X..LMY$ afferma che Y è uno degli L di uno degli M di Y, ossia che X sta nella relazione L con qualche Z che sta nella relazione M con Y; si vede chiaramente che questo non è altro che il prodotto relativo introdotto poi da Peirce. Al solito, $X.LMY$ nega la relazione precedente. $X..lY$ asserisce la relazione contraria di L (dà cioè, in terminologia insiemistica moderna, il complemento della relazione L, che viene in generale indicato da De Morgan

[13] Tra l'altro è proprio nel contesto della generalizzazione della sillogistica che De Morgan enuncia le leggi oggi note appunto col suo nome (e che peraltro sappiamo essere già conosciute dagli Stoici, e comunque chiaramente espresse dai logici medievali) come segue: «Il contrario [la negazione] di un aggregato [di una somma logica] è il composto [prodotto logico] dei contrari degli aggregati; il contrario di un composto è l'aggregato dei contrari dei componenti».

con la lettera minuscola della corrispondente lettera maiuscola che indica la relazione). $X..L^{-1}Y$ dà la relazione conversa di L; se $L = L^{-1}$ la relazione L è simmetrica. Per poter rendere il sillogismo generalizzato, vale a dire con ogni possibile opportuna relazione al posto della copula, occorre poter esprimere la «quantificazione»: così LM' significa un L di *ogni* M, ossia qualcosa che sta nella relazione L con ogni elemento della classe M; mentre $L'\ M$ sta nel senso di «alcuni» (che come caso particolare può anche comprendere tutti o nessuno) vale a dire qualcosa sta nella relazione L con un elemento di M, ammesso che esistano L di questo tipo. Sulla base di queste convenzioni, De Morgan dimostra numerosi teoremi di quella che oggi diciamo teoria delle relazioni; ad esempio fa vedere che il complemento della conversa è la conversa del complemento; che contrarie di converse sono contrarie; che la conversa di un prodotto relativo è il prodotto delle converse invertito; che se una relazione è contenuta (e De Morgan *non* intende ciò in senso insiemistico) in un'altra, allora la conversa della prima è contenuta nella conversa della seconda e altrettanto vale per le contrarie; nonché numerosi altri risultati su cui non è luogo qui insistere.

A questo punto è in grado di fornire la tavola dei sillogismi generalizzati, che egli presenta in due casi diversi, a seconda che la copula L sia la stessa relazione transitiva in entrambe le premesse (i numeri romani stanno per le tradizionali figure):

	1	2	3	4
I	$X..LY$ $Y..LZ$ $X..LZ$	$X.LY$ $Y..L^{-1}Z$ $X.LZ$	$X..LY$ $Y.L^{-1}Z$ $X.L^{-1}Z$	—
II	$X.LY$ $Z..LY$ $X.LZ$	$X..LY$ $Z..L^{-1}Y$ $X..LZ$	$X..LY$ $Z.LY$ $Z.L^{-1}Z$	—
III	$Y..LX$ $Y.LZ$ $X.LZ$	$Y.LX$ $Y..LZ$ $X.L^{-1}Z$	$Y..LX$ $Y..L^{-1}Z$ $X..L^{-1}Z$	—
IV	—	$Y..LX$ $Z.L^{-1}Y$ $X.LZ$	$Y.LX$ $Z..L^{-1}Y$ $X.L^{-1}Z$	$Y..LX$ $Z..LY$ $X..L^{-1}Z$

oppure si considerino nelle due premesse relazioni L e M diverse qualsiasi (col solito significato dei numeri romani):

	1	2	3	4
I	$X..LY$ $Y..MZ$ $X..LMZ$ $X.lM'Z$ $X.L.mZ$ $LM\|N$	$X.LY$ $Y..MZ$ $X..lMZ$ $X.LM'Z$ $X.l,mZ$ $BM^{-1}\|L$	$X..LY$ $Y.MZ$ $X..LmZ$ $X.Lm'Z$ $X.L,MZ$ $L^{-1}N\|M$	$X.LY$ $Y.MZ$ $X..lmZ$ $X.Lm'Z$ $X.L,MZ$ $lm\|N$
II	$X.LY$ $Z.MY$ $X..lM^{-1}Z$ $X.LM^{-1\prime}Z$ $X.l,m^{-1}Z$ $NM\|L$	$X..LY$ $Z..MY$ $X..LM^{-1}Z$ $X.lM^{-1\prime}Z$ $X.L,m^{-1}Z$ $LM^{-1}\|N$	$X..LY$ $Z.MY$ $X..Lm^{-1}Z$ $X.lm^{-1\prime}Z$ $X.L,M^{-1}Z$ $L^{-1}N\|M^{-1}$	$X.LY$ $Z.MY$ $X..lm^{-1}Z$ $X.Lm^{-1\prime}Z$ $X.l,M^{-1}Z$ $lm^{-1}\|N$
III	$Y..LX$ $Y.MZ$ $X..L^{-1}mZ$ $X.l^{-1}m'Z$ $X.L,^{-1}MZ$ $LN\|M$	$Y.LX$ $Y..MZ$ $X..l^{-1}MZ$ $X.L^{-1}M'Z$ $X.l,^{-1}mZ$ $NM^{-1}\|L^{-1}$	$Y..LX$ $Y..MZ$ $X..L^{-1}MZ$ $X.l^{-1}M'Z$ $X.L,^{-1}mZ$ $L^{-1}M\|N$	$Y.LX$ $Y.MZ$ $X..l^{-1}mZ$ $X.L^{-1}m'Z$ $X.l,^{-1}MZ$ $l^{-1}m\|N$
IV	$Y.LX$ $Z.MY$ $X..l^{-1}m^{-1}Z$ $X.L^{-1}m^{-1\prime}Z$ $X.l,^{-1}M^{-1}Z$ $l^{-1}m^{-1}\|N$	$Y..LX$ $Z.MY$ $X..L^{-1}m^{-1}Z$ $X..l^{-1}m-^{1\prime}Z$ $X.L,^{-1}M^{-1}Z$ $LN\|M^{-1}$	$Y.LX$ $Z..MY$ $X..l^{-1}M^{-1}Z$ $X..L^{-1}M^{-1\prime}Z$ $X.l,^{-1}m^{-1}Z$ $NM\|L^{-1}$	$Y..LX$ $Z..MY$ $X..L^{-1}M^{-1}Z$ $X..l^{-1}M^{-1}Z$ $X.L^{-1}M^{-1\prime}-Z$ $L^{-1}M^{-1}\|N$

Per concludere, malgrado alcuni risultati di De Morgan siano stati ripresi come abbiamo già accennato e da Boole e da Peirce, l'ancoramento del logico inglese al modello aristotelico agì in definitiva più come momento bloccante che come stimolo per la sua visione della logica, sicché egli non seppe concretare operativamente possibili prospettive di un calcolo più generale in senso moderno, prospettive che qua e là affiorano, a livello di anticipazione, nei suoi scritti; in particolare per quanto riguarda la teoria delle relazioni non apprezzò in tutta la sua portata il fatto che la teoria stessa, indipendentemente dal suo rapporto con la sillogistica,

costituiva una dottrina autonomamente importante e interessante ogni campo della matematica. Intravvide insomma varie generalizzazioni e aperture possibili, nell'ambito però di un contesto che ormai aveva fatto il suo tempo ed era stato di fatto superato dalla prospettiva booleana.

4. LA «RIVOLUZIONE» BOOLEANA

In che cosa dunque consiste la natura rivoluzionaria dell'impostazione della logica operata da Boole, nella quale culmina quel processo di individuazione del ruolo del «formale» indipendente da interpretazioni privilegiate o comunque precostituite? La risposta è che superando lo sfondo aritmetico sulla base del quale lavoravano gli algebristi inglesi, e al di là della «chiusura» di De Morgan costituita dal privilegiato modello aristotelico, Boole[14] afferma in modo netto e deciso la natura *formale* del calcolo in generale, nel senso che anche l'istituzione di un calcolo logico è una costruzione formale passibile di *diverse* interpretazioni che non costituiscono più la base esclusiva e primaria da cui la struttura formale viene astratta.

[14] George Boole nacque a Lincoln nel novembre del 1815, da famiglia di condizioni assai modeste. Avviato a studi commerciali, si applicò come autodidatta allo studio del latino, del greco, del francese, del tedesco e dell'italiano. Costretto all'insegnamento elementare da necessità familiari, sempre da autodidatta affrontò lo studio della matematica, nel qual campo condivise le idee della scuola di Cambridge anche in conseguenza della sua amicizia personale con un esponente di questa, Duncan Farquharson Gregory; conobbe anche De Morgan di cui divenne amico e col quale ebbe uno scambio epistolare particolarmente interessante. Iniziò la sua attività scientifica con un lavoro del 1841 *Researches on the theory of analytical transformations* (*Ricerche sulla teoria delle trasformazioni analitiche*) cui seguì, nel 1844, *A general method in Analysis* (*Un metodo generale in Analisi*), direttamente collegato ai temi della scuola di Cambridge, in quanto imperniato sul principio della separazione dei simboli. Nel 1847 pubblica *The mathematical analysis of logic; being an essay toward a calculus of deductive reasoning* (*L'analisi matematica della logica: saggio di un calcolo del ragionamento deduttivo*) che è appunto considerata unanimemente come il «manifesto» originario della moderna logica formale. L'anno successivo riespose brevemente l'opera in questione nell'articolo *The calculus of logic* (*Il calcolo della logica*). Nel 1849 viene chiamato come professore di matematica al Queen's College di Cork, ove insegna fino alla morte, avvenuta nel dicembre del 1864. Fra la produzione scientifica di Boole emerge il volume *An investigation of the laws of thought, on which are founded the mathematical theories of logic*

Questo atteggiamento, mentre ci sembra giustifichi appieno l'importanza da noi data nei capitoli precedenti alle varie fasi della evoluzione della scuola algebrica inglese, riconnette direttamente Boole a Leibniz per l'ampiezza dell'orizzonte prospettato; assai illuminanti ci sembrano in proposito le parole di Boole, a partire dalla motivazione stessa dell'*Analysis* del 1847 quando afferma, nella prefazione all'opera: «Nella primavera di quest'anno la mia attenzione fu attratta dalla disputa allora sorta fra Sir W. Hamilton e il professor De Morgan; e fui indotto dall'interesse che la ispirava a riesumare trame, ormai quasi dimenticate, di indagini precedenti. Mi sembrava che, malgrado la logica possa essere riguardata con riferimento all'idea di quantità, essa fosse caratterizzata anche da un altro e più profondo sistema di relazioni. Se era legittimo riguardarla dall'*esterno* come una scienza che attraverso la mediazione del Numero si connette con le intuizioni di spazio e tempo, era legittimo anche riguardarla dall'*interno* come basata su fatti di ordine diverso che hanno la loro sede nella costituzione della mente».

«Coloro che hanno familiarità con lo stato attuale della teoria dell'algebra simbolica – ribadisce Boole nell'introduzione dell'opera – sono consapevoli che la validità dei procedimenti dell'Analisi non dipende dall'interpretazione dei simboli che vi sono impiegati, ma soltanto dalle leggi che regolano la loro combinazione. Ogni sistema di interpretazione che non modifichi la verità delle relazioni che si suppone esistano tra tali simboli è ugualmente ammissibile, ed è così che il medesimo processo può, secondo uno

and probabilities (*Una ricerca sulle leggi del pensiero sulle quali sono fondate le teorie matematiche della logica e delle probabilità*) del 1854, ove Boole, che lo considerava il proprio capolavoro, dà la versione «matura» delle sue idee sulla logica nell'ambito di un preciso contesto filosofico generale; paradossalmente la storiografia logica ha tuttavia considerato quest'opera da molti punti di vista come una pura e semplice «complicazione filosofica» rispetto all'opera del '47. Vanno ancora ricordati *On the application of the theories of probability to the question of the combination of testimonies or judgements* (*Sull'applicazione della teoria della probabilità alla questione della combinazione di testimoni o giudizi*) del 1857 e, del 1862, *On the theory of probability* (*Sulla teoria della probabilità*) che sono entrambi saggi di interpretazione logica del calcolo della probabilità. È del 1859 un *Treatise of differential equations* (*Trattato di equazioni differenziali*) e del 1860 *Calculus of finite differences* (*Calcolo delle differenze finite*). Boole lasciò inoltre parecchio materiale inedito che la sua vedova, signora Mary Everest, donò alla Royal Society; sottoposti gli scritti al giudizio di De Morgan, questi decise che tale materiale nulla avrebbe aggiunto ai risultati già ottenuti e resi pubblici dallo stesso Boole e bocciò quindi il progetto di pubblicarli.

schema di interpretazione, rappresentare la soluzione di una questione riguardante le proprietà dei numeri, secondo un altro schema quella di un problema di geometria e, secondo un altro ancora, quella di un problema di dinamica o di ottica. Questo principio possiede un'importanza fondamentale e si può affermare che i recenti progressi dell'Analisi pura sono stati in larga misura promossi dall'influenza che esso ha esercitato nel dirigere l'indirizzo della ricerca».

Boole può così decisamente affermare che «la caratteristica che definisce un calcolo autentico consiste in questo: che esso è un metodo fondato sull'impiego di simboli le cui leggi di combinazione sono note e generali, e i cui risultati ammettono un'interpretazione coerente. Il fatto che alle forme oggi esistenti di Analisi venga assegnata un'interpretazione quantitativa è il risultato delle circostanze che determinarono il sorgere di tali forme, e noi non dobbiamo farne una condizione universale dell'Analisi. Sulla base di questo principio generale, io intendo appunto fondare il calcolo logico, e reclamare per esso un posto tra le forme di analisi matematica ormai generalmente riconosciute, senza tener conto del fatto che, dati il suo soggetto e gli strumenti di cui si avvale, esso deve, per il momento, rimanere isolato». Si ricordino le parole di Boole che abbiamo riportato nel paragrafo 1 – e che egli scriveva nelle *Laws* del '54 – relative alle applicazioni della logica e alla natura matematica dei suoi processi ultimi, aggiungendovi la conclusione importantissima cui Boole perviene nelle stesse pagine, che «*non fa parte dell'essenza della matematica di essere intimamente connessa con le idee di numero e di quantità*» [corsivo nostro] e si avrà un'idea precisa non solo del senso del «matematismo» booleano, ma anche della nuova originale e moderna concezione che egli ha della matematica cui riferisce il calcolo logico.

Accanto a questo matematismo tuttavia, dalle parole precedenti risulta anche in evidenza un secondo motivo di fondo della prospettiva booleana, cui di fatto sarà legato lo scopo fondamentale di tutta la sua opera e che sarà esplicitamente e ripetutamente affrontato nelle *Laws* del '54: il motivo psicologistico che inquadra tutto il disegno di Boole nella ricerca delle leggi del pensiero al fine di ricavarne informazioni e aperture verso la conoscenza della costituzione della mente. Le leggi del pensiero cioè rifletterebbero la struttura della mente e il metterle in evidenza nei loro nessi «al-

gebrici» può appunto farci compiere un importante passo avanti nella sua conoscenza. Che questo del resto fosse l'effettivo scopo «in grande» delle ricerche, e anzi si può dire globalmente dell'attività scientifica di Boole, è espresso esplicitamente nelle *Laws* quando egli si preoccupa inizialmente di stabilire che «la presente opera non è una ripubblicazione di un precedente trattato dell'Autore, intitolato "L'analisi matematica della logica". La sua prima parte è in effetti dedicata allo stesso oggetto, ed essa comincia con lo stabilire lo stesso sistema di leggi fondamentali, ma i suoi metodi sono più generali, e il suo campo di applicazione molto più vasto. Essa presenta i risultati maturati in alcuni anni di studio e di riflessione, di un principio di indagine connesso alle operazioni intellettuali, la cui prima esposizione venne scritta in poche settimane dopo che ne avevo concepito l'idea»; passa quindi a rendere esplicito l'orizzonte più ampio nel quale situa l'intera sua opera, che ha lo scopo di «investigare le leggi fondamentali di quelle operazioni della mente mediante le quali si effettua il ragionamento; di dare ad esse espressione nel linguaggio simbolico di un calcolo e di stabilire, su questi fondamenti, la scienza della Logica e costruire i suoi metodi; di rendere questo stesso metodo la base di un metodo generale per l'applicazione della teoria matematica delle probabilità; e, infine, di raccogliere dai vari elementi di verità portati alla luce nel corso di questa indagine alcune probabili indicazioni concernenti la natura e la costituzione della mente umana».

La cosa importante da notare è qui, al di là della particolare concezione sostenuta da Boole (e della fortuna della concezione stessa), il suo atteggiamento di apertura per quanto riguarda l'inquadramento del discorso logico e la sua ampia visione del problema generale della conoscenza. Quando Boole dichiara, sempre nelle *Laws*, che «... il solo oggetto della logica non è quello di renderci capaci di dedurre inferenze corrette da date premesse; né il solo scopo del calcolo delle probabilità quello di permetterci di stabilire su basi sicure le questioni di assicurazioni sulla vita... Entrambi questi studi hanno anche un interesse di altro tipo, che loro deriva dalla luce che essi gettano sulle capacità intellettuali. Essi ci istruiscono relativamente al modo col quale il linguaggio e il numero servono come aiuto strumentale al processo di ragionamento; essi ci rivelano in certo grado la connessione fra differenti

capacità dell'intelletto comune; essi ci mostrano quali siano, nei due domini della conoscenza probabile e dimostrativa, gli elementi essenziali della verità e della correttezza – elementi non derivati da alcunché, ma profondamente fondati nella costituzione delle facoltà umane», ci sembra che egli riesca a superare la particolare concezione di cui è sostenitore, o meglio, a rendersi indipendente da essa per lanciare un programma che bandisca la chiusura della ricerca scientifica fine a se stessa; programma o raccomandazione quanto mai attuale anche ai nostri giorni, ove imperativo è ormai divenuto l'impegno di superare la specializzazione particolare del campo di ricerca proprio di ogni singolo studioso, per aprire la mente a una visione globale d'inquadramento umano, politico e culturale (come quello prospettato da Boole stesso).

È così giustificata appieno l'affermazione booleana circa l'originalità e la maggior maturità delle *Laws* rispetto all'*Analysis*: le prime hanno portato a esposizione ampia e approfondita il contesto generale in cui si muove il lavoro e la scoperta tecnica, individuando tale contesto in un rapporto logica-psicologia che finisce col problematizzare la stessa attività simbolizzatrice dell'uomo. Abbiamo già accennato al fatto che in generale la valutazione che la critica ha dato delle *Laws* ha quasi unanimemente trascurato questo momento di maturità, limitandosi a considerarle come un miglioramento tecnico per quanto riguarda l'esposizione del calcolo, senza nemmeno tentare, in generale, un'analisi delle motivazioni più profonde che abbiamo cercato di mettere in luce come moventi effettivi dell'opera di Boole, quando addirittura non è giunta a qualificarle, e di solito su basi puramente pragmatiche, come qualcosa di sovrapposto alla stringatezza e alla nitidezza del sistema simbolico.

Ancora una volta, una particolare sensibilità ha dimostrato a questo proposito Barone nella sua opera già tanto citata nel corso di questi capitoli. A quest'opera rimandiamo il lettore desideroso di approfondire la questione, mentre ci accingiamo a dare un sommario resoconto del sistema logico-algebrico booleano e delle sue interpretazioni in termini di classi, di proposizioni e probabilistiche. Non ci atterremo a una esposizione determinata ma ci riferiremo indifferentemente, talora senza farlo esplicitamente rimarcare, alla presentazione delle *Laws* o a quella dell'*Analysis*.

Supponiamo di partire con un «universo del discorso» ossia

con un insieme di «cose» concrete o no, che Boole indica con 1, per ragioni che saranno subito chiare, e che costituisce per così dire il vero «oggetto» del nostro sistema.[15] Nell'ambito di questo universo potremo compiere degli «atti di elezione», potremo cioè scegliere quegli oggetti x che godano di una proprietà X. Boole chiama i simboli minuscoli x, y, z, u, v, w, ... «simboli elettivi» ed essi stanno tanto per l'atto mentale di «elezione» quanto per il risultato di questo atto, ossia per la classe degli individui che nell'universo godono delle rispettive proprietà X, Y, Z, U, V, W, ... Assunta quindi un'infinità di simboli elettivi, è chiaro che può eseguirsi nell'universo quell'operazione che consiste nel fare due successivi atti di scelta, ossia nello scegliere ad esempio dall'universo prima la classe x degli individui che godono della proprietà X e quindi nello scegliere fra questi ultimi la classe y degli individui che godono della proprietà Y. Indicheremo con $x \times y$ o semplicemente con xy, l'atto o il risultato di tali atti successivi di elezione. È chiaro che si giungerà allo stesso risultato sia scegliendo prima x e poi y, sia invece scegliendo prima y dall'universo 1, e quindi x. Vale a dire si può assumere a livello linguistico che per l'operazione di «*prodotto logico*» sopra illustrata valga la legge

$$xy = yx$$

dove il segno « = » sta a indicare uguaglianza in estensione, indica cioè che le due classi scritte a destra e a sinistra del simbolo stesso contengono esattamente gli stessi elementi. La relazione di «uguaglianza in estensione» è l'unica relazione fra classi che considereremo e di conseguenza « = » sarà l'unico segno relazionale

[15] Si noti che nell'*Analysis* Boole sembra concepire l'universo del discorso come onnicomprensivo, intendendo che «... esso comprende ogni classe concepibile di oggetti, sia che questi esistano, sia che non esistano...». Nelle *Laws* invece, dove il simbolo 1 viene introdotto sulla base di analogie algebriche che vedremo più avanti nel testo, sembra invece intenderlo come vero e proprio «universo del discorso di una teoria» ossia come insieme (classe) degli oggetti di cui «si parla». Analoghe considerazioni possono farsi dualmente per il simbolo 0 che nell'*Analysis* viene impiegato senza che se ne dia una definizione esplicita, mentre nelle *Laws*, sulla base delle analogie algebriche cui sopra si accennava, viene introdotto come una classe particolare, la cui interpretazione logica è il «Nulla» (classe vuota). È opportuno qui aggiungere che il concetto booleano di universo del discorso venne recepito dagli studiosi posteriori nella sua prima accezione onnicomprensiva. Schröder in particolare criticherà questa determinazione mostrando come essa possa portare a contraddizioni.

del sistema. È chiaro inoltre che si ha $1x = x$ qualunque sia la classe x e che se si suppone di avere la «classe vuota», il «Niente», si ha analogamente $0x = 0$ qualunque sia la classe x; in altri termini, lo 0, il niente, è quella classe (unica) che rimane inalterata per qualsiasi atto di elezione in essa.

Analogamente, possiamo pensare all'atto di «elezione» dall'universo della classe degli elementi che sono caratterizzati dal godere della proprietà X o della proprietà Y *ma non di entrambe*, vale a dire dall'appartenere alla classe x o alla classe y, *ma non ad entrambe*. Nel caso allora che le classi x e y siano *disgiunte*, non abbiano cioè elementi in comune, si può pensare definita fra esse un'operazione che verrà indicata con « + » e che applicata alle classi x e y darà appunto la classe $x + y$ degli elementi dell'universo che appartengono ad una almeno delle due classi x e y e a una sola di esse.

Il fatto di aver definito l'operazione « + » di «aggregazione di parti in un tutto» o, come diremo, di «*somma logica*», per classi disgiunte, è un'assunzione caratteristica del sistema booleano, cui probabilmente Boole si decide per potersi assicurare l'introduzione della differenza $x - y$ fra classi (facilmente interpretabile da parte del lettore sulla scorta delle considerazioni precedenti) in modo tale che le due operazioni « + » e « – » risultino nel sistema l'una inversa dell'altra. D'altra parte questa determinazione non permette un comportamento simmetrico dell'operazione di somma logica rispetto a quella di prodotto logico e comporta altre difficoltà che appariranno in seguito, sicché è stata in generale riguardata dai continuatori di Boole come uno dei più immediati difetti nel suo sistema.

Nel 1986 T. Hailperin (in *Boole's Logic and Probability* [*La logica di Boole e la probabilità*]) ha indagato più da vicino le ragioni di questa scelta booleana di definire la somma solo fra classi disgiunte e ha mostrato come sullo sfondo si possa porre un modello di «algebra delle classi» diverso da quello per noi oggi usuale, nel quale ha senso contare il numero delle volte in cui un elemento (o meglio, una sua copia) compare in una collezione. La struttura che emerge è quella di una «heap algebra» (algebra dei mucchi), come la definisce Hailperin, che non coincide con la usuale algebra di Boole che tradizionalmente si assume come sfondo dei suoi lavori. L'introduzione di queste algebre, secondo Hailperin, fu motivata dal desiderio di mantenere il parallelismo fra somma logica di due proposizioni e somma delle probabilità di due even-

ti (che è la somma aritmetica solo quando gli eventi sono indipendenti). Questo mostrerebbe come sin dall'inizio Boole mirasse ad un'analisi simultanea di inferenza deduttiva e di inferenza probabilistica.

Assunte quindi come fondamentali nel suo sistema le operazioni «×», «+», «−» e come relazione la «=», Boole definisce il «complemento» $1-x$ di un termine x come quella classe di elementi dell'universo che non appartengono a x; e stabilisce quella che è la legge caratteristica del sistema, la cosiddetta «legge degli indici», che si esprime nella forma $x^2 = x$ (o, ma solo nell'*Analysis*, nella forma più generale $x^n = x$; nelle *Laws* non assume invece questa forma generale per evitare talune difficoltà sulle quali è superfluo qui insistere). Sulla base delle operazioni e relazioni precedenti, alcune delle quali abbiamo sopra illustrato, si può dire che Boole assume, fra implicite ed esplicite, le seguenti «leggi del pensiero» per i simboli «logici» o «elettivi» del suo sistema

1. $xy = yx$	Commutatività del prodotto
2. $x + y = y + x$	Commutatività della somma
3. $z(x + y) = zx + zy$	Distributività del prodotto rispetto alla somma
4. $z(x - y) = zx - zy$	Distributività del prodotto rispetto alla differenza
5. Se $x = y$, allora:	
$zx = zy$ $z + x = z + y$ $x - z = y - z$	Sostitutività di «=» rispetto alle operazioni «×», «+» e «−»

6. $x^2 = x$ o equivalentemente $x(x - 1) = 0$ Legge degli indici.

Si noti che le leggi (gli «assiomi» potremmo dire) 1)-5) sono tutte applicabili anche nell'ordinaria algebra numerica, nel senso cioè che si trasformano in proposizioni vere ove ai segni di operazione e relazione si dia il significato ordinario e si interpretino i segni, x, y, z, come numeri; la legge distintiva del sistema booleano è la 6). Qui il sistema sembra scostarsi irrimediabilmente dai modelli numerici e d'altronde la 6) si dimostrerà di fondamentale importanza per ridurre le formule del calcolo in forma conveniente all'interpretazione logica. Si noti tuttavia che la legge in questione è soddisfatta dai numeri 1 e 0, ed è proprio a questo punto che Boole traduce operativamente il motivo conduttore della sua ri-

cerca, affermando che «invece di determinare la misura dell'accordo formale dei simboli della logica con quelli dei Numeri in generale, è più immediatamente suggerito di confrontarli con i simboli di quantità *che ammettono solo i valori* 0 e 1». Supposto quindi di considerare un'algebra i cui simboli ammettano come valori solo 0 e 1, allora «le leggi, gli assiomi e i processi di tale algebra saranno identici nella loro estensione completa alle leggi, gli assiomi e i processi di un'algebra della logica. Esse sarebbero divise solo da differenze di interpretazione. E il metodo di quest'opera è fondato su questo principio».

Si noti che questo non significa affatto che il sistema di Boole sia un'algebra a due valori; per poter affermare ciò dovrebbe trovarsi, fra le leggi sopra elencate, una proposizione quale «$x = 0$ o $x = 1$» (o della forma equivalente «Se $x \neq 0$ allora $x = 1$») che invece non vi figura. Quello che Boole afferma è semplicemente che il suo sistema è *interpretabile* in termini delle quantità numeriche 0 e 1, ossia che questi due numeri costituiscono per così dire un insieme di «cose» che rende vere tutte e sei le leggi poste a fondamento del proprio sistema (con adeguata interpretazione, è ovvio, dei segni di operazione e relazione). Del resto, nelle nostre argomentazioni precedenti a scopo esemplificativo abbiamo implicitamente assunto che il sistema di Boole fosse interpretabile in termini di classi: e il lettore può convincersi facilmente che questa interpretazione non soddisfa la condizione di «bivalenza» che abbiamo sopra espresso, dal momento che ovviamente non tutte le classi di un dato sistema di classi soddisfano in generale la condizione di essere uguali all'universo (classe totale) o al niente (classe vuota).

Stabilito il «sommario» delle leggi fondamentali del pensiero che debbono reggere il processo inferenziale, Boole individua come segue le condizioni che debbono essere soddisfatte per condurre un ragionamento corretto con l'aiuto dei simboli; allo scopo è necessario che «1) ai simboli venga assegnata un'interpretazione fissata nell'espressione dei dati; e che le leggi di combinazione di questi simboli siano correttamente determinate da quell'interpretazione; 2) che i processi formali di soluzione o dimostrazione siano condotti sempre in obbedienza a tutte le leggi sopra determinate, senza riferimento alla questione dell'interpretazione dei particolari risultati ottenuti; 3) che il risultato finale sia interpretabile

in forma e che esso sia effettivamente interpretabile in accordo con quel sistema di interpretazione che era stato impiegato nell'espressione dei dati».

Questo passo delle *Laws* è estremamente significativo in quanto si può dire riassuma la concezione globale di Boole relativamente allo *status* del calcolo logico. È chiaro infatti che Boole propone, nell'ambito di una determinazione contenutistica del discorso inferenziale, una conduzione puramente formale, in questo senso per lui matematica, del discorso stesso. In termini moderni potremmo esprimere la cosa dicendo che Boole individua e distingue i due momenti semantico e sintattico e pur non riguardandoli in assoluto come separati o autonomi l'uno rispetto all'altro, privilegia il momento sintattico per quanto riguarda la conduzione dell'argomentazione; l'interpretazione semantica interviene nel momento iniziale in cui per così dire si fissano i principi formali del calcolo, e nel momento finale di questo, quando cioè si interpretano i risultati. È proprio questo momento centrale di totale riferimento «algebrico» il momento matematizzante «negativo» nel pensiero booleano, a parere di alcuni suoi critici posteriori, segnatamente di Jevons; ed è innegabile che questa completa analogia algebrica di fondo ha consentito a Boole passaggi e risultati intermedi non immediatamente interpretabili in senso logico, e la cui giustificazione finale non è fra le pagine più lucide della sua costruzione.

Per giustificare la possibilità di passi intermedi nella deduzione di equazioni logiche non immediatamente interpretabili logicamente, Boole, nel capitolo quinto delle *Leggi*, fa appello ad un principio generale secondo lui non ulteriormente giustificabile, che sta al fondo della stessa possibilità del conoscere. È lo stesso principio che per Boole giustifica nei «calcoli trigonometrici» il ricorso al simbolo $\sqrt{-1}$ che non ha interpretazione. Il principio fondamentale del ragionamento simbolico cui Boole fa riferimento sembrerebbe in sostanza affermare che – poiché ragionando simbolicamente si usano *solo* regole formali – se queste hanno un modello che pure non è conforme all'interpretazione che abbiamo in mente, la conclusione del ragionamento simbolico varrà anche per la nostra interpretazione. Nel nostro caso è l'interpretazione in $\{0,1\}$ a garantire la coerenza delle regole del calcolo. È difficile valutare la portata del principio (che ricorda quello di permanenza di Hankel) ma è importante sottolineare come qui Boole, anche se

in modo approssimativo, individui un aspetto centrale del procedere simbolico – la sua indipendenza da specifiche interpretazioni – che dovrebbe giustificare la validità di una sorta di principio di transfer: se una deduzione è valida in una interpretazione lo sarà in ogni altra che soddisfa le regole del calcolo. Questa possibilità di passare da un'interpretazione a un'altra che si presenta *solo* se le interpretazioni considerate hanno determinati rapporti che Boole non esplicita, è per Boole la ragione della fecondità e l'essenza stessa del procedere formale, intuizione questa che verrà rigettata dai suoi immediati successori. Ma, va ribadito, è con questa sua aderenza anche se talora apparentemente eccessiva ad analogie algebriche che Boole riesce concretamente, e per la prima volta nella storia della logica, ad andare al di là e a superare decisamente il modello sillogistico aristotelico: proprio per l'effettiva generalità intrinseca del calcolo proposto, il sillogismo risulterà un caso particolarissimo di un più generale e comprensivo schema inferenziale, di più generali metodi di «ragionamento» la cui controparte algebrica verrà appunto stabilita mediante il *calculus ratiocinator* booleano.[16] Prima di passare a esporre brevemente i

[16] Queste considerazioni ci permettono di trattare in nota la questione relativa alla traduzione e alla «resa» nel linguaggio booleano di quello che è il sistema classico della sillogistica. Cominciamo col notare che non simbolizzando Boole la relazione di inclusione fra classi, ne viene subito che le proposizioni (in particolare le categoriche) saranno rappresentate da equazioni nelle quali «qualcosa» è posto uguale a 0 o a 1 (essendo come abbiamo visto l'identità l'unica relazione per cui nel linguaggio si disponga di un apposito simbolo) oppure ricorrendo alla quantificazione del predicato. Boole in effetti impiega entrambi i metodi per la traduzione delle categoriche e ciò costituisce un ulteriore «neo» della sua sistemazione. È chiaro infatti che le categoriche potrebbero tradursi come segue. Per le generali

A $x(1-y)=0$ Ogni x è y, ossia è vuota la parte comune a x e al complemento di y.

E $xy=0$ Nessun x è y, ossia è vuota la parte comune a x e y.

Per quanto riguarda le particolari, le cose non sono così semplici. Sarebbe immediato renderle mediante disuguaglianze, ad esempio come segue

I $xy \neq 0$ Qualche x è y ossia non è vuota la parte comune a x e y.

O $x(1-y) \neq 0$ Qualche x non è y, ossia...

Ma come detto sopra Boole ricorre in questo caso alla quantificazione del predicato, introducendo un particolare simbolo elettivo v che in qualche modo è destinato a rappresentare «alcuni» o «qualche» e traduce le due proposizioni in questione come segue

I $xy = v$

O $x(1-y) = v$

momenti fondamentali di tale calcolo, vogliamo osservare che il collegamento qui implicitamente fatto con Leibniz sarà uno dei punti più pregnanti nelle obiezioni che, come vedremo, Frege solleverà alla sistemazione booleana, proprio per difendersi dall'accusa di alcuni suoi contemporanei che gli imputavano di non essersi adeguato al modello in questione.

Il processo fondamentale che regge tutto il calcolo booleano è quello dello *sviluppo* di una funzione. Si cominci con l'osservare che una qualunque espressione del calcolo contenente il simbolo elettivo x può essere riguardata come una funzione di x, e può essere rappresentata in forma generale abbreviata col simbolo $f(x)$; analogamente si potrà parlare di funzioni $f(x, y)$, $f(x, y, z)$,... di due, tre,... argomenti. Ora per quelle funzioni in cui x, y, z,... sono *simboli logici*, o simboli di quantità suscettibili di assumere i soli valori 0 e 1, in breve simboli che soddisfano la legge 6) degli indici (o di «dualità» come verrà anche chiamata nelle *Laws*) è possibile dare uno *sviluppo standard* nella forma

$$f(x) = ax + b(1 - x)$$

con la caratterizzazione puramente negativa, per quanto riguarda la classe rappresentata da v, che essa deve essere considerata «indefinita sotto ogni rispetto salvo il fatto di contenere almeno un elemento» (vale a dire può essere una qualunque classe non vuota). Questo espediente non è certo fra i più felici e potrebbe portare a risultati palesemente contraddittori (si pensi semplicemente che dalle equazioni $x = v, y = v$ si potrebbe dedurre $x = y$ per le due classi x, y qualunque non vuote; o analogamente da $xy = v$ e $zt = v$ si giungerebbe a $xy = zt$, e così via), risultati che Boole riesce sì ad evitare sempre ma, si può dire, per sua personale abilità piuttosto che in forza della struttura del calcolo. Con l'introduzione comunque di un secondo termine w analogo a v si ha la seguente rappresentazione delle categoriche

A	$x = vy$	$x(1 - y) = 0$
E	$x = v(1 - y)$	$xy = 0$
I	$vx = w(1 - y)$	$v = xy$
O	$vx = w(1 - y)$	$v = x(1 - y)$

Comunque, con questa simbolizzazione Boole può rendere tutta la teoria del sillogismo classico, riportando la teoria delle inferenze immediate (conversioni, obversioni, contrapposizioni, ecc.) o a elementari proprietà dei simboli elettivi o a semplici trasformazioni di tipo *algebrico* sulle equazioni che traducono le proposizioni; è analogamente chiaro che la «conclusione» di un sillogismo da due premesse è così riportata, in termini per ora intuitivi, all'eliminazione dalle premesse stesse del simbolo elettivo che rappresenta il termine medio, vale a dire a stabilire con metodi algebrici dalle date equazioni delle premesse una terza opportuna equazione che non contenga il termine medio.

dove a e b sono determinati in modo tale da rendere il risultato equivalente (dal punto di vista dell'interpretazione logica, o in senso numerico per i soli valori 0 e 1) alla funzione data.[17]

Premesso che con $f(1)$, rispettivamente $f(0)$ si indica il valore della funzione $f(x)$ nella quale a x si sia sostituito 1 o 0 rispettivamente, il richiesto sviluppo per una funzione di una variabile sarà

$$f(x) = f(1)x + f(0)\,(1 - x).$$

Ad esempio, se $f(x) = \dfrac{1 + x}{1 + 2x}$
avendosi $f(1) = 2/3$ e $f(0) = 1$, si otterrà

$$\frac{1+x}{1+2x} = \frac{2}{3}\,x + (1 - x)$$

e, nel senso precedentemente chiarito, questa identità vale per ogni possibile valore di x. Analogamente si procede per funzioni di due, tre,... variabili, per le quali si ha, rispettivamente:

$$\begin{aligned} f(x,y) &= f(1,1)xy + f(1,0)x(1-y) + f(0,1)\,(1-x)\,y + \\ &\quad + f(0,0)(1-x)(1-y); \\ f(x,y,z) &= f(1,1,1)xyz + f(1,1,0)xy(1-z) + f(1,0,1)x(1-y)z + \\ &\quad + f(0,1,1)(1-x)yz + f(1,0,0)x(1-y)\,(1-z) + \\ &\quad + f(0,1,0)(1-x)y(1-z) + + f(0,0,1)(1-x)(1-y)z + \\ &\quad + f(0,0,0)(1-x)(1-y)(1-z) \end{aligned}$$

e così via per un numero superiore di variabili; da questi esempi dovrebbe risultar facile al lettore ricavare il semplice metodo che permette di generalizzare il procedimento a funzioni con un numero (finito) qualunque di variabili.

[17] Si noti che mentre nell'*Analysis* Boole giunge a dare l'espressione per lo sviluppo di una funzione in modo spregiudicato e largamente non giustificato a partire dall'analogo sviluppo di Taylor di una funzione in serie di potenze, nelle *Laws* invece riduce a una nota questa analogia, ricorrendo però anche in questo caso a motivazioni analogiche di tipo algebrico, che, a rigore, non risultano comprensibili e giustificabili appieno se non si tiene conto delle argomentazioni dell'*Analysis*.

In una forma sviluppata gli elementi $f(1), f(0); f(1,1), f(1,0),\ldots;$ $f(1,1,1),\ldots$ si dicono *coefficienti* (o *moduli*); gli elementi x, $(1 - x)$; xy, $(1 - x)y$, $x(1 - y),\ldots;$ xyz, $(1 - x)yz,\ldots$ si dicono invece *costituenti*; gli elementi di quelli fra i costituenti che non contengono complementi (ossia termini della forma $(1 - x)$ ecc.), vengono detti *fattori*: così, x, y, z, sono ad esempio i fattori del costituente xyz. Fondamentali sono le tre seguenti proprietà dei costituenti di qualsiasi sviluppo:

1) ogni costituente t soddisfa la legge di dualità, ossia $t(t - 1) = 0$;
2) la somma di tutti i costituenti di ogni sviluppo è 1;
3) il prodotto di due qualunque costituenti di un dato sviluppo è 0.

Boole rimarca che tali proprietà valgono *sempre* e «sono comprensibili» per simboli quantitativi che assumano i valori 0 e 1 soltanto e osserva che esse possono ritenersi soddisfatte anche «per simboli logici, *sempre se interpretabili*» e che malgrado non necessariamente una *funzione* risulti interpretabile (si intenda sempre: *logicamente* interpretabile) dopo lo sviluppo, tuttavia le *equazioni* godono di detta proprietà e sono sempre riducibili mediante questo processo a forme interpretabili. In altri termini, il problema è ora quello di dare un significato *logico* e ai costituenti di uno sviluppo e ai rispettivi coefficienti, e in particolare si tratta di vedere come questi ultimi influiscano sul significato, sulla «pregnanza» dei primi.

Per quanto riguarda i costituenti si vede subito che essi sono sempre interpretabili, dal momento che rappresentano «le diverse divisioni esclusive dell'universo del discorso, formate dalla predicazione e negazione in ogni possibile modo delle qualità denotate dai simboli $x, y, z,\ldots$ ecc.»; si può allora passare all'interpretazione di una data equazione mediante la regola seguente: si sviluppi la funzione corrispondente e si uguagli a 0 ogni costituente il cui coefficiente non è nullo;[18] l'interpretazione complessiva dei risul-

[18] I costituenti i cui coefficienti sono nulli vengono ovviamente eliminati dall'espressione dello sviluppo di una funzione, nel senso che essi indicano una sorta di indipendenza logica della funzione sviluppata dalle particolari «combinazioni» dei termini che seguono i coefficienti in questione.

tati così ottenuti darà l'interpretazione della data equazione. Per chiarire la cosa prendiamo un esempio dallo stesso Boole; si consideri la definizione di «animale puro» della legge ebraica: «sono animali puri quelli che sono contemporaneamente ruminanti e hanno il piede caprino», e poniamo che sia:

x = animali puri
y = animali dal piede caprino
z = animali ruminanti.

La data proposizione può allora essere espressa dall'equazione $x = yz$, che si riduce immediatamente alla forma $x - yz = 0$. Sviluppandola col metodo sopra esposto si ottiene:

$$0xyz + xy(1-z) + x(1-y)z + x(1-y)(1-z) + (1-x)yz + \\ + 0(1-x)y(1-z) + 0(1-x)(1-y)z + 0(1-x)(1-y)(1-z).$$

Trascurando per quanto detto i costituenti i cui coefficienti sono uguali a 0, e uguagliando a 0 i costituenti rimasti, si ha

$$xy(1-z) = 0, \; xz(1-y) = 0, \; x(1-y)(1-z) = 0, \; (1-x)yz = 0$$

e ognuna di queste condizioni ci dà per così dire una parziale informazione logica relativa all'intero contenuto dell'equazione da cui siamo partiti. Dalla prima ad esempio ricaviamo che non «esistono» (in base alla definizione precedente, di «puro») animali puri, dal piede caprino, ma non ruminanti; la seconda ci informa che non esiste la classe degli animali puri, ruminanti, ma non aventi il piede caprino (ossia che questa classe è vuota); la terza che non esiste la classe degli animali puri non ruminanti e senza piede caprino, la quarta infine che non esiste la classe dei ruminanti col piede caprino che non siano puri. Come si vede la somma di queste informazioni parziali esaurisce il significato dell'informazione contenuta nell'equazione di partenza; in questo senso affermavamo prima l'equivalenza del risultato dello sviluppo con l'equazione data.[19]

[19] È interessante riportare la conclusione cui giunge Boole alla fine di quest'esempio. Egli infatti ne deduce che ogni proposizione primaria (si veda avanti nel

Il vero problema interpretativo che si presenta a Boole per quanto riguarda il procedimento di sviluppo di una funzione si ha quando si pensi di avere a che fare con funzioni affatto generali, che non necessariamente siano poste nella forma $V = 0$ o $V = 1$, ma abbiamo invece la forma generale $V = w$, dove V è una funzione dei simboli logici x, y, z, ecc. e w è un qualunque simbolo logico (è chiaro che i precedenti sono casi particolari di questo). In generale infatti potranno presentarsi delle forme i cui coefficienti sono non interpretabili. Limitiamoci a chiarire la cosa con un esempio dello stesso Boole, che ci permetterà nel contempo di renderci conto di un secondo processo del calcolo booleano, quello di *soluzione* di un'equazione: esso consiste nel determinare un elemento di una proposizione elettiva, per quanto complessa, come funzione logica degli altri elementi dell'equazione stessa.

Supponiamo di avere l'equazione (che esprime, per una certa interpretazione dei simboli, una proposizione)

$$x = zy.$$

Risolvere quest'equazione, ad esempio rispetto a z, significa, con totale analogia algebrica, ricavare il valore di z in funzione degli altri termini della proposizione; si ha cioè

$$z = \frac{x}{y}.$$

Il risultato così ottenuto non è però interpretabile, in quanto non è stata introdotta nel sistema di Boole l'operazione di divisione. Qui si coglie la fondamentalità del processo di sviluppo: invece infatti di *eseguire* quell'operazione, limitiamoci, come abbiamo fatto sopra, ad *esprimerla*; quindi sviluppiamo l'espressione così ottenuta. Nel nostro caso si ha:

$$z = xy + \frac{1}{0}x(1-y) + 0(1-x)y + \frac{0}{0}(1-x)(1-y)$$

testo) può essere «risolta in una serie di negazioni di esistenza di certe definite classi di cose e, da questo sistema di negazioni, può essere a sua volta ricostruita». Al problema, che sorge immediatamente, di come si possa passare da una serie di negazioni ad una affermazione positiva, egli risponde che «...la mente assume l'esistenza dell'universo non *a priori* come fatto indipendente dall'esperienza, ma *a posteriori* come deduzione dall'esperienza, o *ipoteticamente* come fondamento della possibilità del ragionamento assertorio».

e tutto si riduce al problema di dare un'interpretazione a coefficienti quali 1/0 o 0/0, o, più in generale, a coefficienti numerici qualsiasi.

Boole determina questa interpretazione, sulla base di analogie algebriche che non è qui il caso di descrivere ma che comunque non risultano certo sempre convincenti. Potrà essere utile al lettore immaginare la questione in termini di espressione di un certo contenuto logico relazionale fra i termini di una data equazione che, una volta eseguito lo sviluppo, si traduce in «gradi di presenza» o in «assenza» di certi costituenti in quello che potremmo chiamare lo «spazio logico» della proposizione data. Ciò premesso, «1) Il simbolo 1 come coefficiente di un termine dello sviluppo indica che va presa la totalità della classe che quel costituente rappresenta. 2) Il coefficiente 0 indica che quella classe non va considerata. 3) Il coefficiente 0/0 indica una porzione del tutto *indefinita* della classe, ossia che vanno presi *alcuni*, *tutti* o *nessuno* dei suoi membri. 4) Ogni altro simbolo (quindi in particolare 1/0 e ogni altro valore numerico diverso da 0 e da 1 come coefficiente), indica che il costituente cui esso è prefisso va uguagliato a zero».[20]

Altro processo fondamentale del calcolo booleano è quello dell'*eliminazione* di un termine x da una data equazione. La cosa si basa sul fatto che se $f(x) = 0$ è un'equazione logica, che comprende il simbolo elettivo x, allora l'equazione

$$f(0)f(1) = 0$$

è vera indipendentemente dall'interpretazione di x: questa equazione rappresenta appunto il risultato dell'eliminazione di x dalla

[20] In scritti inediti posteriori alle *Laws*, Boole identifica i quattro coefficienti 1,0, 0/0, 1/0 con quattro categorie di pensiero e precisamente con la categoria dell'universale, del non esistente, dell'indefinito e dell'impossibile, che costituiscono forme necessarie del ragionamento e che Boole confronta con le categorie kantiane trovandole superiori e più adeguate di queste ultime perché quelle da lui proposte «hanno una relazione reale e scientifica e non puramente immaginata con le possibilità del pensiero logico. Non c'è alcuna parte», prosegue Boole, «delle speculazioni di Kant che a prima vista sia più stringente della sua tavola sistematica delle categorie; essa è tuttavia l'apparenza dell'ordine scientifico senza la realtà». Vedremo come Jevons riuscirà a dare a tali coefficienti un'interpretazione puramente logica.

equazione $f(x) = 0$.[21] Per eliminare un simbolo da una data equazione si applicherà quindi la seguente regola: si portano tutti i termini al primo membro (cambiando ovviamente di segno) si dà al simbolo da eliminare il valore 1 e quindi il valore 0; si moltiplicano i risultati delle due sostituzioni effettuate e si uguaglia il prodotto a zero. Si voglia ad esempio eliminare il simbolo v, e interpretare il risultato ottenuto, dall'equazione

$$y = vx$$

che, intendendo per y «uomo» e per x «mortale», è come sappiamo una possibile traduzione simbolica della proposizione universale «tutti gli uomini sono mortali». Si avrà successivamente:

$$\begin{array}{rr} & y - vx = 0 \\ \text{per } v = 1 & y - x = 0 \\ \text{per } v = 0 & y = 0. \end{array}$$

Moltiplicando i risultati e uguagliando a zero: $y(1 - x) = 0$ ossia, ovviamente: «non esistono uomini non mortali». Si noti che pur essendo il risultato ottenuto già interpretabile direttamente, si può proseguire per maggiori «informazioni» sviluppando la funzione

$$1 - x = \frac{0}{y}$$

e ottenendo quindi

$$1 - x = \frac{0}{0}(1 - y)$$

cioè, ricordando quanto sopra detto sulle interpretazioni dei coefficienti: «coloro che non sono mortali non sono uomini» vale a dire la contrapposizione dell'equazione data. Consideriamo un altro esempio. La proposizione «nessun uomo (y) è perfetto (x)» viene resa, come sappiamo, da

$$y = v(1 - x)$$

[21] La dipendenza di questo processo da quello di sviluppo di una funzione (a conferma appunto del ruolo assolutamente fondamentale di quest'ultimo) appare evidente nella dimostrazione che Boole dà – e che noi per semplicità non riportiamo – di questo asserto e della regola che su di esso si basa.

Eliminiamo v. Scritto $y - v(1-x) = 0$ e posto $v = 1$ si ha: $y - (1-x) = 0$; posto ora $v = 0$ si ottiene $y = 0$, sicché il risultato **dell'eliminazione** sarà $y - y(1-x) = 0$, ossia $yx = 0$, che sappiamo già interpretare («non esistono uomini perfetti»). Sviluppando ora la funzione si ottiene, risolvendo rispetto ad x

$$x = \frac{0}{y} = \frac{0}{0}(1-y)$$

vale a dire «nessun essere perfetto è uomo». Analogamente, sviluppando $1 - x$ si avrebbe

$$1-x = 1 - \frac{0}{y} = \frac{y}{y} = y + \frac{0}{0}(1-y)$$

ossia: «gli esseri imperfetti sono tutti gli uomini più una porzione indeterminata (tutti, alcuni, nessuno) di altri esseri che non sono uomini».

Ulteriore processo introdotto da Boole è quello di *riduzione*, che consiste nel ricavare da una data serie di (due o più) equazioni una sola equazione la quale conservi tutte le «informazioni» relative ai rapporti logici fra i termini contenuti nella serie di equazioni iniziali; in altre parole, che risulti equivalente al sistema di equazioni dato. Anche per questo processo la regola è molto semplice e viene come al solito dimostrata sulla base di proprietà dello sviluppo delle singole funzioni del sistema: da un numero qualunque di equazioni della forma $V = 0$ con V che soddisfa la legge di dualità $V(V - 1) = 0$ (brevemente, equazioni elettive o logiche) si può ottenerne una sola combinando le equazioni date per semplice somma e tenendo conto del fatto che ogni altra equazione della forma $V = 0$ può essere ridotta, elevandola al quadrato, a un'equazione cui sia applicabile la regola in questione.

A questo punto è chiaro che mentre la soluzione delle equazioni elettive rappresenta una generalizzazione del classico processo di inferenza immediata (e ne abbiamo mostrato alcuni esempi) sempre sulla base dello sviluppo si ha una generalizzazione effettiva anche del processo di inferenza mediata, in particolare del sillogismo, che risulta in questo contesto essere un caso del tutto speciale di applicazione dei processi di riduzione ed eliminazione: precisamente

quello in cui si tratta di ridurre *due* equazioni a una sola e di eliminare *un solo* termine, il termine medio. Ne viene immediatamente che l'estensione del metodo booleano – comprendendo l'esplicitazione dei rapporti logici (le «inferenze») fra un numero (finito) qualsiasi di proposizioni in un numero (finito) qualsiasi di termini – prospetta effettivamente un calcolo del tutto generale, che si è decisamente staccato dalle «strettoie» del modello aristotelico.

Il significato logico del calcolo booleano risulta però in modo compiuto solo una volta che si analizzano ulteriormente le sue diverse interpretazioni. Nelle pagine precedenti abbiamo parlato – anche se in modo non sistematico – di almeno due interpretazioni possibili, alle quali abbiamo fatto costantemente ricorso nelle esemplificazioni: quella in termini di classi e quella in termini di valori numerici 0 e 1; abbiamo anche notato che queste due interpretazioni, fra l'altro, differiscono in particolare per quanto riguarda l'eventuale presunta «bivalenza» del sistema di Boole: mentre infatti la prima *non* soddisfa la disgiunzione esclusiva «$x = 0$ o $x = 1$», è immediatamente ovvio che tale proposizione è soddisfatta dalla seconda interpretazione. Non sono solo queste le interpretazioni possibili e le due per noi più interessanti sono quella proposizionale e quella probabilistica.

Per quanto riguarda la prima, occorre innanzitutto fare una precisazione. Tutte le proposizioni logiche possono considerarsi, secondo Boole, come appartenenti a una di due classi esaustive e disgiunte: quella delle proposizioni primarie o concrete e quella delle proposizioni secondarie o astratte. Le proposizioni della prima classe esprimono asserzioni su fatti o eventi, sulla loro mutua dipendenza, parlano in altri termini di *cose*; le proposizioni della seconda classe parlano invece delle possibili relazioni che questi eventi hanno fra loro, rispetto alla loro eventuale verità o falsità; in altri termini, parlano di *proposizioni*. Questa distinzione, afferma Boole, è in pratica molto vicina, anche se non coestensiva, alla «comune distinzione logica fra proposizioni categoriche e proposizioni ipotetiche». Assumendo questa distinzione, il discorso precedente, in termini di classe (o sillogismi) si riferiva a proposizioni categoriche o primarie; l'interpretazione di cui ora vogliamo parlare è in termini di proposizioni secondarie o ipotetiche, che sono costituite da due o più categoriche unite da una copula o congiunzione, e che traggono i loro vari nomi (condizionali, disgiuntive,

ecc.) appunto dal tipo di congiunzione usata. L'interpretazione in termini di proposizioni secondarie (diremo semplicemente: proposizioni, ove ciò non dia luogo a confusioni) riguarda allora proposizioni non analizzate di cui si studiano quei rapporti inferenziali che non dipendono dalla struttura delle proposizioni stesse, ma dal loro essere vere o false in dipendenza delle «congiunzioni» che le originano.

Ora, per Boole calcolo delle classi e calcolo delle proposizioni hanno una stessa struttura formale e vengono differenziati solo a livello delle interpretazioni ed è quella un'ulteriore prova della maggior profondità del suo calcolo rispetto all'analisi tradizionale. È però possibile ricondurre la seconda alla prima, interpretando il simbolo 1 come l'universo ipotetico che «comprenderà tutti i casi e le congiunture di circostanze concepibili». I simboli elettivi $x, y, z, \ldots$ rappresenteranno corrispondentemente la scelta di tutti i casi in cui le proposizioni X, Y, $Z, \ldots$ sono vere. Per una data proposizione Y si hanno le due sole possibilità che essa sia vera o falsa e poiché «questi casi presi assieme esauriscono l'universo» delle proposizioni e il primo caso è definito dal simbolo elettivo y, il secondo sarà definito dal simbolo $1 - y$. Nel caso invece si considerino più proposizioni, esse daranno luogo a una data combinazione di casi «il cui numero dipenderà dal numero delle considerazioni estranee che abbiamo introdotto».

Per due proposizioni X e Y si avrà ad esempio la situazione seguente, che traiamo dall'*Analysis*.

Casi	Espressioni elettive
1) X vera, Y vera	xy
2) X vera, Y falsa	$x(1 - y)$
3) X falsa, Y vera	$(1 - x)y$
4) X falsa, Y falsa	$(1 - x)(1 - y)$.

Si osservi che in questo caso la somma delle espressioni elettive che tengono conto di tutti i casi possibili è 1; orbene, la cosa si può generalizzare considerando che «l'estensione dell'universo ipotetico non dipende affatto dal numero delle circostanze prese in considerazione... e che per quanto grande possa essere il numero di quelle circostanze, la somma delle espressioni elettive rappresentanti ogni caso concepibile sarà sempre l'unità».

Nei passi dell'*Analysis* sopra riportati appare chiara un'interpretazione in termini proposizionali con le «congiunzioni» (i connettivi, diremmo noi oggi) fra proposizioni intese estensionalmente come funzioni di verità. Nelle *Laws* tuttavia questa prospettiva muta bruscamente, perché Boole riduce la considerazione delle «circostanze» alla sola dimensione temporale, cosicché il simbolo 1 rappresenterà «l'universo o la totalità del tempo a cui si suppone il discorso si riferisca in qualche modo» e una proposizione del tipo ad esempio: «Se *X* è vera allora anche *Y* è vera» dovrà essere intesa come «il *tempo* in cui la proposizione *X* è vera è il *tempo* in cui la proposizione *Y* è vera.» A parte questa «deviazione» booleana delle *Laws*,[22] il sistema formale si presta in entrambi i casi, a parere di Boole, a una pressoché completa elaborazione di quelle forme di inferenza oggi analizzate entro la «logica delle proposizioni» (o degli enunciati).

Veniamo così a quella che può considerarsi l'«interpretazione probabilistica» del sistema di Boole e che a partire almeno dal 1848 costituì una delle motivazioni centrali di Boole per lo sviluppo della sua teoria e sicuramente un elemento nuovo rispetto alla logica tradizionale, ripreso dalla maggior parte dei continuatori di Boole. Accennato a questa possibilità alla fine dell'*Analysis*, Boole dedica vari capitoli delle *Laws* al problema, su cui riferiremo tuttavia molto brevemente. Il punto di fondo è che Boole afferma esplicitamente che «la teoria generale e il metodo della logica ... costituiscono anche la base di una teoria e del corrispondente metodo delle probabilità». Ecco perché è stato prima necessario provvedere a una sistemazione della logica e «la ragione di questa necessità di un metodo precedente per la logica come base della teoria delle probabilità può essere espressa in poche parole. Prima di poter determinare il modo in cui la frequenza attesa dell'occorrenza di un particolare evento dipende dalle frequenze note dell'occorrenza di altri eventi qualsiasi, dobbiamo essere a conoscenza delle mutue dipendenze di questi eventi. Parlando tecnicamente, dobbiamo essere in grado di esprimere l'evento di cui si richiede la probabilità in funzione degli eventi di cui le probabilità sono

[22] Va ricordato peraltro che questa da noi detta «deviazione» viene riguardata col dovuto interesse dagli odierni cultori della cosiddetta «logica del tempo» (*tense logic*). Si veda ad esempio l'appendice A del volume *Time and modality* (*Tempo e modalità*) di Arthur N. Prior (Oxford 1957, 1968^2).

date. Ora questa determinazione esplicita appartiene in tutti i suoi esempi al dipartimento della logica». L'inferenza probabilistica è quindi un'inferenza di tipo logico e così tanto il ragionamento deduttivo che quello probabilistico rivelano la loro base comune.

Noi ci limiteremo qui ad osservare che in certo senso quanto fa Boole nelle *Laws* equivale sostanzialmente a dare una interpretazione del calcolo in termini di teoria della probabilità (intesa in senso classico, ossia come rapporto fra i casi favorevoli e i casi possibili) in cui i simboli elettivi vengano cioè sostituiti da lettere che stanno per probabilità di un dato evento, secondo la tavola

Eventi		Probabilità
1) xy	occorrenza di x e y	pq
2) $x(1-y)$	occorrenza di x senza y	$p(1-q)$
3) $(1-x)y$	occorrenza di y senza x	$(1-p)q$
4) $(1-x)(1-y)$	assenza congiunta di x e y	$(1-p)(1-q)$

Tanto l'interpretazione in termini di proposizioni secondarie quanto quella in termini probabilistici *non* soddisfano la condizione che abbiamo chiamato di «bivalenza» del calcolo di Boole. Il punto fondamentale è che in questo modo Boole mostrava come la struttura algebrica sottogiacente al calcolo della probabilità fosse quella del calcolo delle classi anticipando in un senso quello che A.N. Kolmogorov avrebbe fatto nel 1933 fondando la teoria della probabilità su quella della misura e aprendo dall'altro la strada a quella concezione logicista della probabilità vista come legame logico fra enunciati che avrebbe trovato negli anni trenta di questo secolo in John Maynard Keynes e in Rudolf Carnap gli assertori più convinti.

5. L'ALGEBRA DELLA LOGICA NELL'OTTOCENTO DOPO BOOLE

La profonda originalità e l'angolazione del tutto caratteristica e innovatrice con cui Boole aveva impostato la teoria dell'inferenza logica erano destinate a trovare immediata risonanza negli am-

bienti scientifici mondiali. L'introduzione booleana dei simboli elettivi e la trattazione generale del processo di sviluppo sono indubbiamente elementi che caratterizzano il suo sistema in quanto «algebrico» ma abbiamo visto che Boole (pur con le cautele e le precisazioni necessarie presentate nelle pagine precedenti, e trascurando l'interpretazione «temporale» delle *Laws*) nell'interpretazione del suo calcolo in termini di proposizioni (secondarie) era giunto assai vicino – come potremmo dire in termini moderni – alla nozione di «valore di verità» di una proposizione e a una considerazione dei connettivi quali «funzioni di verità». In altri termini, aveva embrionalmente scritto le prime «tavole di verità». In certo senso un tipo d'approccio di questo genere sarà caratteristico e si può dire costituirà un punto di partenza per la definitiva sistemazione della logica in forma «non algebrica» o, come l'abbiamo chiamata, «logicista», quale verrà intrapresa da Frege a partire dal 1879.

Quando Frege presenterà la propria sistemazione, i suggerimenti booleani saranno già stati potentemente sviluppati sicché la proposta fregeana verrà in generale riguardata come «contrapposta» a quella booleana, ovviamente ritenuta superiore; vedremo che in effetti nello sviluppo del filone algebrico si giungerà, in particolare con Peirce, a tutta una serie di risultati «propri» della sistemazione logicista e che inoltre alcuni risultati fondamentali per l'odierna ricerca logica (*tout court*) saranno scoperti nel secondo decennio del nostro secolo nell'ambito e nella scia della sistemazione schröderiana del filone algebrico. Torneremo nell'ultimo paragrafo dedicato alla logica algebrica, sulla contrapposizione fra queste due tradizioni; per adesso occupiamoci dello sviluppo dell'algebra della logica dopo Boole.

Come si ricorderà accanto alla messe di suggestioni originali e feconde riscontrabili nell'elaborazione booleana, abbiamo posto in luce anche taluni «nei» che sommariamente riguardavano: la limitazione relativa all'operazione «+», che poteva aver luogo solo fra classi disgiunte; le connesse questioni delle operazioni inverse (di cui in particolare la «divisione» risultava logicamente non interpretabile) e dell'ammissione di coefficienti numerici diversi da 0 e 1 (anch'essi di interpretazione «problematica»); l'introduzione del simbolo v e analoghi.

Nel torno di circa mezzo secolo dall'apparizione delle opere di

Boole tutta una serie di ricercatori si adoperò a eliminare questi «nei» e a isolare la struttura di quella che oggi viene comunemente chiamata «algebra di Boole». Questa sorta di «revisione» ebbe inizio essendo Boole ancora in vita, con la *Pure logic* (*Logica pura*, 1864) di Jevons e per gli immediati successori di Boole si esplicò lungo tre direttrici principali: 1) critica alla concezione esclusiva della somma logica, con conseguente introduzione nel sistema di una somma non esclusiva (Jevons, Peirce, Grassmann, McColl, Schröder; unica voce fedele, almeno in un primo momento, alla concezione booleana della somma, quella di J. Venn); 2) introduzione della possibilità di esprimere nel sistema l'*esistenza*, perché, come afferma Peirce, «la notazione di Boole è solo capace di esprimere che qualche descrizione di cose *non* esiste, e non è in grado di dire che qualcosa *esiste*» (McColl, Peirce) e analisi della relazione di identità «=» (Jevons, Peirce, Schröder); 3) introduzione di nuove notazioni per estendere l'algebra di Boole a un'algebra di relazioni (Ellis, De Morgan, Peirce, Murphy, Macfarlane, Schröder).

Prenderemo in esame nel prossimo paragrafo lo sviluppo puramente algebrico delle indagini di Boole; qui ci limitiamo ad osservare che ben pochi lavori di logica «scientifica» della seconda metà dell'Ottocento si situano in un contesto non booleano: escluso, come già accennato, Frege, si può ancora ricordare in questo senso Robert Grassmann (*Die Begriffslehre oder Logik* [*La teoria del concetto o logica*, 1872]; *Die Logik un die andern logischen Wissenschaften* [*La logica e le altre scienze logiche*, 1890]); Joseph Rémy Léopold Delboeuf,[23] e qualche altro autore; per il resto la gran parte degli scritti di logica formale dell'ultimo periodo dell'Ottocento si riallaccia per un verso o per l'altro a una problematica booleana e abbiamo già accennato al fatto che anche De Morgan presenta qualche lavoro in quest'ambito di idee.

Fra gli altri logici inglesi si possono ricordare, oltre al già citato Jevons, Charles Graves (*Mathematical expressions for hypothetical and*

[23] Ecco come Jevons dà notizia dell'opera di Delboeuf nell'introduzione alla seconda edizione (postuma, 1884) dei suoi *Studies in deductive logic* (*Studi di logica deduttiva*): «Gli scritti di M. Delboeuf sulla logica algoritmica pubblicati per la prima volta nella "Revue Philosophique" del 1876... sono molto interessanti, ma furono scritti nella totale ignoranza da parte dell'Autore di quanto era stato fatto in proposito in questo paese da parte di Boole e di altri».

disjunctive propositions [*Espressioni matematiche per le proposizioni ipotetiche e disgiuntive*, 1850]); Robert L. Ellis[24] (*Notes on Boole's laws of thought* [*Note sulle leggi del pensiero di Boole*, 1863]); Robert Harley (*Remarks on Boole's mathematical analysis of logic* [*Osservazioni sull'analisi matematica della logica di Boole*, 1867]); e ancora, come già detto, Hugh McColl, John Venn, Alexander Macfarlane[25] e molti altri. Fuori dall'Inghilterra ricordiamo, in America, ovviamente Peirce e la sua scuola alla Johns Hopkins University (in particolare: George Bruce Halsted, Christine Ladd Franklin, Oscar Howard Mitchell, A.B. Kempe e altri); in Germania, ovviamente Schröder e ancora Jacob Lüroth, Eugen Müller, Alwin Korselt; in Russia, Platon Sergejevič Poretsky e Nikolai A. Vasiliev; in Italia si riallacciano a Boole, anche se in modo del tutto particolare, Peano (e parte della sua scuola) come pure Alfonso del Re e Albino Nagy; è infine opportuno ricordare il francese Louis Liard con il volume *Les logiciens anglais contemporains* (*I logici inglesi contemporanei*) del 1878.

Dopo questo rapidissimo cenno panoramico che dovrebbe tuttavia dare al lettore la misura della fecondità delle nuove idee booleane, vediamo ora di prendere in esame i più rappresentativi degli autori su nominati, riservandoci fin da ora di dedicare il prossimo paragrafo ai due più significativi, ossia a Peirce e Schröder.

Cominciamo senz'altro da Stanley Jevons[26] che si inserisce

[24] Per quanto riguarda Ellis, Jevons dà notizia di una contenuta polemica sorta fra loro circa la rispettiva «originalità». Nell'opera citata alla nota precedente, Jevons riconosce un argomento per il quale è debitore a Ellis; per il resto si limita a osservare che gli scritti di cui Ellis rivendica l'indipendenza furono pubblicati dopo i suoi lavori che trattano o gli stessi argomenti, o argomenti molto affini. Peirce viceversa annovera Ellis fra gli studiosi che più hanno contribuito all'impostazione della logica delle relazioni.

[25] Sempre da Jevons, *op. cit.*: «Il dottor Macfarlane di Edimburgo ha pubblicato (1879) una nuova versione del sistema di Boole sotto il titolo *Algebra of logic* (*Algebra della logica*) nella quale tuttavia io non sono riuscito a scoprire un qualsiasi miglioramento rispetto a Boole».

[26] William Stanley Jevons (1835-82), allievo di De Morgan, fu logico ed economista, professore di filosofia morale ed economia politica all'Owen College di Manchester fra il 1866 e il 1876; passò quindi fino al 1880 all'University College di Londra quale professore di economia politica. Fra le sue opere logiche ricordiamo: *Pure logic, or the logic of quality apart from quantity: with remarks on Boole's system and on the relation of logic and mathematics* (*Logica pura, ovvero la logica della qualità indipendentemente dalla quantità, con osservazioni al sistema di Boole e sul rapporto fra logica e matematica*) del 1864; *The substitution of similars, the true principle of reasoning, de-*

nella scia booleana come irriducibile avversario di quello che per lui era l'aspetto negativo del «matematismo» di Boole. Jevons si propone di eliminare le «troppo ardite analogie con la matematica, e di fare dell'algebra della logica una "logica pura"» che corrisponda effettivamente alla «logica del pensiero comune» e che quindi è *la* logica, l'unica possibile. L'obiettivo di Jevons ad esempio ne *I principi della scienza* è di elaborare una concezione unitaria dei metodi deduttivi e induttivi, probabilistici e non, della scienza. In questo contesto si situa il problema dei rapporti fra logica e matematica: «Nella mia opinione la logica è la scienza superiore, la base generale della matematica come di tutte le altre scienze. Il numero non è che discriminazione logica e l'algebra una logica altamente sviluppata. Così è facile capire la profonda analogia che Boole ha sottolineato fra le forme di deduzione algebrica e logica».

Jevons già nel suo scambio epistolare con Boole (1863-64) discute l'interpretazione esclusiva che questi aveva dato della somma, e che nel sistema di Boole impedisce per esempio di assumere una legge quale la $x + x = x$ che invece viene considerata da Jevons come una legge «autoevidente del pensiero» e qualificata come «legge dell'unità». Per Jevons le «leggi che espri-

rived from a modification of Aristotle's dictum (*La sostituzione degli identici, il vero principio del ragionamento, derivato da una modificazione del* dictum *di Aristotele*) del 1869; *Elementary lessons on logic deductive and inductive* (*Lezioni elementari di logica deduttiva e induttiva*) del 1870. Nel 1869 Jevons aveva costruito una macchina logica, che presentò alla Royal Society nel 1870 (*On the mechanical performance of logical inference* [*Sulla esecuzione meccanica dell'inferenza logica*, 1870]) di cui esiste una descrizione e un disegno in quella che rappresenta la grande opera metodologica del nostro autore: *The principles of science. A treatise on logic and scientific method* (*I principi della scienza. Trattato di logica e metodo scientifico*) pubblicata in due volumi nel 1874 e ristampata tre anni dopo in un unico volume. Della sua produzione quale economista ricordiamo infine *The theory of political economy* (*La teoria dell'economia politica*, 1871) e *The state in relation to labour* (*Stato e lavoro*, 1881). Val qui la pena di ricordare che nell'Inghilterra del secolo scorso questo collegamento fra logica ed economia non fu sporadico. Numerosi sono gli esempi tratti dall'economia politica in Boole, e di economia politica si occuparono Whately, Whewell, Venn, più tardi i Keynes (padre e figlio) e negli anni trenta F.P. Ramsey. In un senso, se si considera che l'economia politica compariva nelle università inglesi come ramo della scienza morale, questo mostra come la logica fosse vista come una sorta di «matematica» per le scienze morali, da contrapporsi o meno alla matematica *tout court*. Anche in questo caso la matematizzazione della logica veniva a sconvolgere classificazioni e distinzioni canoniche.

mono la vera natura e le condizioni dei poteri discriminanti della mente sono la legge di identità, la legge di dualità o del terzo escluso e la legge di contraddizione», accanto al principio di sostituzione degli identici, che per Jevons è il fondamento di tutti i principi inferenziali. Si noti che – coerentemente al suo assunto di costituire una logica «naturale» – egli afferma di considerare le leggi sulle quali fonderà il proprio sistema in comprensione piuttosto che in estensione ma nonostante le argomentazioni da lui portate in appoggio a questa decisione, in effetti tutto il suo calcolo è esprimibile e interpretabile estensionalmente in termini di classi.

Punto di partenza di Jevons, come per Boole, è l'idea che il problema di fondo della logica sia: «dato un numero qualunque di premesse o condizioni logiche», descrivere «ogni classe di oggetti o ogni termine in quanto dipendenti da tali condizioni». Secondo Jevons in questa indagine il suo sistema è superiore a quello di Boole per tre motivi fondamentali:

«1) Ogni processo è di natura e forza autoevidente ed è governato da leggi altrettanto semplici e primarie degli assiomi di Euclide; 2) il processo è infallibile e non fornisce risultati anomali o non interpretabili; 3) le inferenze possono essere condotte con molto meno lavoro di quello richiesto nel sistema del prof. Boole, dove in generale si richiedono un calcolo e uno sviluppo particolare per ogni inferenza». Questa pretesa superiorità lascia alquanto perplessi, specialmente nel suo terzo punto, vista l'estrema complicazione di alcuni casi di inferenza nel sistema di Jevons.

Trascurando tuttavia questo elemento di forzato confronto, che se da una parte è naturale, dall'altra è indubbiamente frutto di vari malintesi da parte di Jevons, primo fra tutti il fraintendimento di ciò che Boole intendesse con «matematica», a Jevons va tuttavia riconosciuto il merito di aver introdotto alcune modifiche nel sistema booleano che rappresentano decisamente dei miglioramenti e che come tali saranno accettate da (quasi) tutti gli studiosi allora impegnati in quest'opera di «revisione». È quindi conveniente, prima di fare qualche osservazione conclusiva su Jevons, dare i tratti essenziali del sistema da questi costruito; lo descriveremo in termini intensionali (ossia originali) e, per confronto, estensionali (classi). I simboli impiegati sono:

$A, B, C, \ldots$	Qualità o gruppi di qualità che costituiscono la parte comune, o significato intensivo, di termini. Estensionalmente, classi.
$a, b, c \ldots$	Corrispondenti termini negativi (che con De Morgan, da cui prende la notazione, Jevons chiama anche «contrari»).
$A = B$	Il segno $=$ significa «identità di significato dei termini fra i quali è posto; così $A = B$ indica che le qualità significate da A sono identiche alle qualità significate da B». Estensionalmente: A e B sono classi identiche.
A •|• B[27]	Il segno •|• significa «alternazione non esclusiva... Così A •|• B indica le qualità di A o quelle di B» senza escludere che coincidano. Estensionalmente: ciò che è A o B o entrambe.
AB	La giustapposizione di termini dà un termine il cui significato è «la somma delle qualità significate dalle due lettere». Estensionalmente: ciò che è A e B.
0	Ciò che è contraddittorio (o, con parole di Jevons, «ciò che è escluso dal pensiero»). Estensionalmente, come in Boole, la classe vuota o il «Niente».

Le leggi di combinazione sono le seguenti (ci limiteremo a commentare o quelle meno evidenti o quelle più significative, riportando in generale la denominazione jevonsiana)

[27] Jevons assume il simbolo •|• per indicare l'alternativa non esclusiva, per differenziarsi il più possibile, anche a livello simbolico, dalla notazione matematica. Nel seguito, per comodità tipografica, noi useremo tuttavia il solito simbolo « + », che ovviamente non va confuso con l'identico simbolo impiegato nella descrizione del sistema booleano.

1) Se $A=B$ e $B=C$ allora $A=C$	Transitività di « = ».
2) $AB=BA$	Legge di commutazione della giustapposizione.
3) $A(B+C)=AB+BC$	Legge di distribuzione.
4) $AA=A$	Legge di semplicità.
5) $A+A=A$	Legge di unità.
6) $A+B=B+A$	Legge dell'ordine indifferente (commutatività di « + »).
7) $A+0=A$	(In effetti Jevons impiega questa legge senza enunciarla esplicitamente)
8) $A+AB=A$	Legge di assorbimento. Questa legge permette una diretta semplificazione non possibile nel sistema di Boole. Si può estendere a un numero qualunque (finito) di termini.

Jevons chiama le tre leggi successive «il fondamento di ogni ragionamento» e le esprime come segue

9) $A=A$	Legge di identità.
10) $A=AB+Ab$	Legge di dualità o del terzo escluso.
11) $Aa=0$	Legge di contraddizione.

Come si vede Jevons non usa un simbolo per l'universo del discorso (l'1 di Boole) ma in vece sua afferma che «la successiva applicazione della legge di dualità a due, tre, quattro, cinque o più termini origina lo sviluppo di tutte le possibili combinazioni logiche, detto l'*alfabeto logico*...». Per un qualsiasi numero (finito) n di termini A, B, C, ... tale alfabeto consiste, come si verifica facilmente, di 2^n termini che, in linguaggio booleano, altro non sono che i costituenti dello sviluppo di 1. Così, se si hanno due termini A e B,

l'alfabeto logico consisterà dei 4 (2^2) termini AB, Ab, aB, ab; per $n = 3$, analogamente, si avrà ABC, ABc, Abc, aBC, abC, aBc, Abc, abc, e così via; ciò è in effetti equivalente ad assumere l'universo 1.

L'unica regola di inferenza assunta da Jevons è la «regola di sostituzione degli identici: per ogni termine sostituire ciò che in qualche premessa è detto essere identico in significato a quel termine», vale a dire, si può dovunque sostituire b con a se si ha come premessa $b = a$ (questa regola ammette un'ovvia e immediata interpretazione estensionale che anzi è la più «naturale», potendo la su espressa regola di Jevons, col suo riferimento all'identità di significato, condurre in errore). Come De Morgan, Jevons assume che ogni termine debba avere il suo contrario («legge dell'infinito») e questo gli permette di ridurre il numero degli schemi di inferenza. Sulla base di questo linguaggio, Jevons può formulare il suo «criterio di coerenza»: «Due o più proposizioni qualunque sono contraddittorie se e solo se, dopo aver effettuato tutte le possibili sostituzioni, esse originano la scomparsa totale di un qualunque termine, positivo o negativo, dall'alfabeto logico». È sfruttando questo criterio, che in un certo senso fornisce una semantica in cui i modelli sono gli analoghi dei costituenti booleani, che Jevons affronta i problemi di coerenza, equivalenza, deducibilità fra insiemi di proposizioni.

Seguiamo su un esempio l'applicazione del metodo di Jevons all'inferenza, il che ci permetterà di fare i dovuti confronti e rapporti col metodo di sviluppo booleano. Premettiamo che, essendo l'unica relazione del sistema l'identità (« = »), l'espressione delle categoriche classiche nel sistema di Jevons è, sostanzialmente come per Boole, la seguente:

A	Ogni *A* è *B*	$A = AB$
E	Nessun *A* è *B*	$A = Ab$
I	Qualche *A* è *B*	$CA = CAB$
O	Qualche *A* non è *B*	$CA = CAb$

dove il simbolo C svolge qui la stessa funzione del v booleano, ossia sta a indicare «qualche».[28] Consideriamo ora un esempio dello

[28] Ci esimiamo qui dal riprendere le discussioni circa la difficoltà di questa assunzione e dal riferire sulle ampie argomentazioni dedicate da Jevons alla precisazione della problematica ad essa relativa nonché al significato della stessa.

stesso Jevons, rapportabile al caso della «soluzione» booleana di un'equazione. Invece di usare come Boole procedure «algebriche» Jevons sviluppa l'alfabeto logico relativo alla data equazione e quindi effettua tutte le prove (sostituzioni) possibili. Si applichi ad esempio il procedimento all'equazione $A = BC$, e si voglia trovare un'espressione per il termine b in dipendenza della data equazione (la nostra «premessa»). L'equazione contiene tre termini e il suo alfabeto logico sarà pertanto:

$$ABC,\ ABc,\ AbC,\ aBC,\ Abc,\ aBc,\ abC,\ abc,$$

in quanto si tratta di fare ogni possibile combinazione dei tre termini e dei loro negativi. Se ora combiniamo ognuna di queste espressioni con entrambi i membri della premessa, otteniamo, tenendo conto delle leggi del sistema:

	Primo membro (A) della premessa	Secondo membro (BC) della premessa
1)	ABC	ABC
2)	ABc	$ABCc$
3)	AbC	$ABbC$
4)	Abc	$AbBCc$
5)	$AaBC$	aBC
6)	$AaBc$	$aBcC$
7)	$AabC$	$abBC$
8)	$Aabc$	$aBbCc$

Ora Jevons afferma che, se una qualsiasi delle combinazioni dell'alfabeto logico costituisce una contraddizione (ossia dà luogo a almeno un termine della forma Aa) con *un solo* membro della premessa, tale elemento viene detto *combinazione contraddittoria* o soggetto contraddittorio; se una combinazione dell'alfabeto logico è contraddittoria con *entrambi* i membri della premessa la si dirà (combinazione o) *soggetto escluso*; se infine una di tali combinazioni non contraddice *nessuno* dei due membri della premessa verrà detto *soggetto incluso*. Un soggetto incluso o escluso è un soggetto *possibile*; un soggetto contraddittorio è un soggetto *impossibile.* Nel nostro esempio si vede immediatamente che sono contraddit-

torie in questo senso le combinazioni 2), 3), 4), 5) dell'alfabeto logico, mentre le rimanenti sono combinazioni possibili, e precisamente la 1) rappresenta un soggetto incluso, le 6), 7), 8) soggetti esclusi.

La richiesta espressione per il termine b si otterrà ora come segue: cancellate le combinazioni contraddittorie, si cerchino fra le rimanenti (soggetti possibili) quello che contengono il termine b; esse sono la 7) abC, e la 8) abc. La richiesta espressione per b sarà allora $b = abc + abC$. Si tratta di una procedura che ricorda gli odierni metodi di decisione fondati sulle forme normali disgiuntive e si comprende come Jevons tendesse a meccanizzare un metodo del tipo sopra esposto giungendo appunto alla costruzione di una «macchina logica». Il metodo sopra presentato è estremamente simile a quello dei diagrammi di Venn cui accenneremo tra breve e d'altra parte – proprio in quanto riconducibile a quelle che noi oggi chiamiamo forme normali disgiuntive – è perfettamente leggibile in termini booleani: l'alfabeto logico consiste infatti di quei termini la cui somma è 1, ossia costituisce come dicevamo l'universo in senso booleano; alle varie combinazioni 1) soggetto contraddittorio, 2) soggetto incluso, 3) soggetto escluso, corrispondono termini che nello sviluppo booleano hanno coefficienti rispettivamente 0 o 1/0, 1, 0/0 o v. Ricordando inoltre che nell'analogo caso booleano i termini con coefficiente di cui a 1) vengono uguagliati a zero, si comprende come sostanzialmente i due procedimenti siano identificabili. Come osserva Lewis, «nel suo complesso i metodi di Jevons corrono facilmente il rischio di essere noiosi e non posseggono certo nitore matematico. Supponiamo ad esempio di avere tre equazioni contenenti in tutto sei termini. L'"alfabeto logico" consisterebbe di 64 combinazioni, ognuna delle quali andrebbe separatamente provata per ogni equazione, dando così luogo a 192 operazioni separate. Jevons ha enfatizzato la sua differenza da Boole al punto di rifiutare molto di quanto avrebbe fatto meglio a mantenere».

Ciò malgrado, a parte cioè la sua visione dogmatica della logica «naturale» e le pretese semplificazioni al metodo booleano, che abbiamo visto essere più apparenti che reali, a Jevons va riconosciuto il merito di aver posto in atto precisi miglioramenti del sistema di Boole che si sono rivelati molto fecondi per gli sviluppi successivi.

Si tratta in particolare, riassumendo:

a) dell'eliminazione delle operazioni inverse di sottrazione e divisione dal sistema di Boole;

b) della interpretazione della somma $A + B$ in senso non esclusivo. Ciò comporta in particolare l'eliminazione di coefficienti numerici diversi da 0 e 1, sicché oltre a permettere una «naturale» simmetria nell'espressione delle leggi, costituisce una delle modifiche più significative;

c) della introduzione della legge di assorbimento fra i principi dell'algebra booleana.

Pienamente conscio del potere generalizzante del formalismo e quindi non in opposizione all'atteggiamento booleano, ma in implicita polemica con le rigide concezioni di Jevons, si pone il logico inglese John Venn (1834-1923) lettore di logica e scienze morali all'università di Cambridge. Venn presenta le sue idee nelle due edizioni di *Symbolic logic* (*Logica simbolica*, prima edizione 1881, seconda 1894)[29] che Barone definisce come «il primo documento di uno schietto interesse storico per le vicende della logica formale: testimonianza importante del fatto che la riflessione filosofica sugli effettivi nuovi sviluppi scientifici della disciplina non erano affatto necessariamente determinati in un'unica direzione». Venn propone come proprio programma di ricerca «l'esame della logica simbolica, ossia le sue relazioni con la logica comune e il pensiero e il linguaggio comuni: la determinazione e spiegazione di ogni regola ed espressione simbolica generale sulla base di principi puramente logici, piuttosto che semplicemente sulle loro giustificazioni formali; l'invenzione e l'impiego di un nuovo schema di notazione

[29] In quest'opera Venn – al quale risale l'introduzione del termine «logica simbolica» – raccoglie e sistematizza sue precedenti pubblicazioni sull'argomento, fra le quali vanno almeno ricordate: *Consistency and real inference* (*Coerenza e inferenza reale*); *Boole's logical system* (*Il sistema logico di Boole*) entrambe del 1876; *On the diagrammatic and mechanical representations of propositions and reasoning* (*Sulle rappresentazioni diagrammatiche e meccaniche delle proposizioni e del ragionamento*) del 1880; *On the form of logical propositions* (*Sulla forma delle proposizioni logiche*) dello stesso anno e, ancora, i due saggi *On the various notations adopted for expressing the common propositions of logic* (*Sulle varie notazioni adottate per esprimere le comuni proposizioni della logica*) e *On the employment of geometrical diagrams for the sensible representation of logical propositions* (*Sull'impiego di diagrammi geometrici per la rappresentazione concreta delle proposizioni logiche*), pubblicati su «Mind» fra il 1880 e il 1883. Nella prima edizione della *Symbolic logic* Venn non accetta di modificare la definizione di somma logica data da Boole; si adeguerà all'uso ormai invalso nella seconda edizione dell'opera.

diagrammatica che risulterà in perfetta armonia con le nostre generalizzazioni», dichiarando inoltre esplicitamente che il suo lavoro non vuol essere «in alcun modo un commento o una critica a Boole».

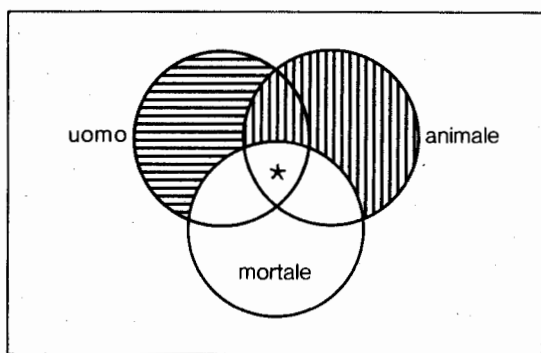
D'altra parte, Venn considera «la logica simbolica e la matematica come branche di un linguaggio di simboli che posseggono alcune, anche se assai poche, leggi di combinazione in comune. Questa comunità di legislazione o uso, nella misura in cui esiste, è la nostra principale giustificazione per l'adozione di un sistema uniforme di simboli per entrambe». Non si tratta quindi di una rigida contrapposizione fra logica e matematica (col predominio implicito dell'una sull'altra) né d'altra parte di una sostituzione della logica «comune» con la logica formale; sicché quest'ultima non è una schematizzazione «necessaria» del modo di ragionare comune, bensì solo uno «dei diversi modi con cui si può ragionare». Dopo aver messo in evidenza questa anticipazione – da parte di Venn – di una sorta di «principio di tolleranza» in logica, passiamo brevemente a considerare le sue proposte originali più propriamente «tecniche», ricordando innanzitutto la sua determinazione, che si differenzia nettamente dall'uso tradizionale, di simbolizzare le proposizioni universali tenendo conto del fatto fondamentale che ogni proposizione universale può essere posta in forma negativa: è chiaro infatti che ad esempio invece di «ogni uomo è mortale», è equivalente dire «nessun uomo è non mortale». Con questo punto di partenza, simboleggiate opportunamente le universali (nell'esempio precedente, con x = uomo e y = mortale, si ha per le categoriche A: $x\bar{y} = 0$; e analogamente per le E: $xy = 0$) e considerati i rapporti logici delle universali con le particolari, si giunge per queste ultime alla simbolizzazione seguente: per le I: $xy > 0$; per le O: $x\bar{y} > 0$. Avendo così posto l'accento sulla portata (carattere) esistenziale delle particolari, non è più lecita la conversione *per accidens* e la subalternazione, sicché si riduce il numero dei sillogismi validi. Questo punto di vista verrà assunto anche da Schröder nella sua sistemazione dell'algebra di Boole e costituisce un primo esempio di quella attenzione alla logica su domini non necessariamente non vuoti, che è alla base della recente logica libera.

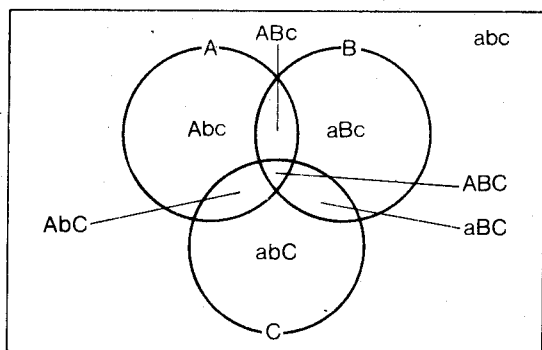
L'altro apporto originale di Venn consiste, come già accennato, nell'introduzione di una rappresentazione diagrammatica per la

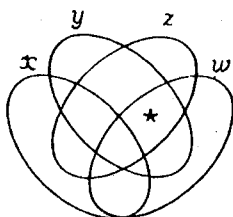
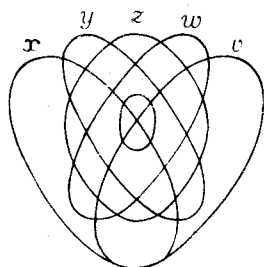
conduzione e la verifica delle inferenze. Come Venn stesso chiarisce, il suo metodo si differenzia da quello euleriano (i «cerchi» di Eulero) perché fornisce una specifica «quantità di informazione» circa la portata delle premesse di un ragionamento, permettendone la specifica e completa conclusione. Si tratta di servirsi di regioni del piano fra loro sovrapponentesi per rappresentare relazioni fra classi o condizioni di verità fra proposizioni.

Caratteristico del metodo di Venn è che ogni possibile combinazione viene rappresentata da un'area diversa (e qui sta la stretta analogia con «l'alfabeto logico» di Jevons o con lo «sviluppo» di Boole); quindi in ogni regione viene segnata, sulla base delle premesse, l'informazione relativa all'appartenenza o meno di elementi alla regione stessa (vale a dire se la classe è vuota o no nel caso delle classi, se la proposizione è vera o no nel caso delle proposizioni). Facciamo un semplice esempio col sillogismo

Ogni uomo è animale
Ogni animale è mortale
―――――――――――
Ogni uomo è mortale

cui corrisponde il seguente «diagramma di Venn»

nel quale: il rettangolo rappresenta l'universo del discorso (è chiaro che non è necessario tracciarlo effettivamente); le aree tratteggiate indicano che le corrispondenti classi sono vuote (e ciò si ricava dalle premesse, con la traduzione «negativa» cui sopra si accennava: non esistono uomini non animali, né animali non morta-

li); l'asterisco in un'area indica invece che la corrispondente classe non è vuota e dà la conclusione del sillogismo del nostro esempio. Se si volesse condurre con questo metodo il problema di «risoluzione» prima visto col metodo di Jevons, a partire dall'equazione $A = BC$, si avrebbe il seguente diagramma:

da cui risulta evidente la scomposizione dell'«alfabeto logico» e dal quale il lettore, tenuto conto della equazione data, può facilmente trovare un'espressione per *b* (che si troverà ovviamente coincidente con quella di Jevons).

Il metodo di Venn è teoricamente estendibile a un numero (finito) qualsiasi di termini; inoltre, almeno in linea di principio, ci si potrebbe servire di qualunque tipo di linea chiusa nel piano per isolare le regioni. Tenuto però conto che ogni «combinazione» o «costituente» deve essere rappresentato da una sua particolare regione, si vede subito che si possono usare cerchi fino a che si tratti con due o tre termini; essi non sono più idonei quando si dovesse trattare con quattro termini (16 combinazioni possibili): in questo caso Venn propone di servirsi di ellissi come mostrato in figura

e, nel caso di cinque termini, propone la seguente figura composita che evita il ricorso a curve di base più complicate

Un metodo diagrammatico analogo a quello proposto da Venn venne impiegato da Charles Lutwidge Dodgson (Lewis Carroll) per la trattazione del sillogismo nella sua *Symbolic logic* (*Logica simbolica*) del 1896.

Altro autore estremamente interessante di questo periodo, momento di mediazione fra le ricerche «algebriche» e quelle «logiciste», è l'inglese Hugh McColl (1837-1909) il quale, con motivazioni di tipo probabilistico, affronta la questione delle difficoltà dello sviluppo della logica delle proposizioni indipendentemente da quella delle classi. Avevamo visto, in proposito, come in Boole il punto di partenza fosse per così dire rovesciato rispetto a quello che oggi si considera come usuale e normale: nella costruzione assiomatica del calcolo delle classi viene in qualche modo già presupposta la «disponibilità» di un'argomentazione di tipo proposizionale (si pensi semplicemente alla stessa enunciazione di alcune leggi del calcolo booleano nella forma «Se... allora...»). È stato anche ripetutamente detto che l'ordine oggi accettato di «dipendenza», secondo il quale, per così dire, l'elemento minimale del ragionamento formale è la logica delle proposizioni (o degli enunciati, come anche diremo) è stato stabilito in modo sistematico da Frege nel 1879. Si può ben dire però che già nel 1872 il problema veniva affrontato nella sua specificità e, almeno in senso parziale e senza raggiungere la sistematicità e organicità della versione fregeana, veniva risolto appunto da McColl: questi parte come dicevamo da una problematica connessa al calcolo delle probabilità e approda alla conclusione che il momento logico fondamentale debba essere quello enunciativo, dal momen-

to che le proposizioni risultavano il mezzo linguistico più idoneo a rappresentare gli «eventi».

Dalle prime intuizioni del 1872, McColl viene precisando il suo disegno, e viene esponendo i propri risultati in tutta una serie di lavori che vanno dai sette articoli *The calculus of equivalents statements and integration limits* (*Il calcolo di enunciati equivalenti e limiti d'integrazione*), pubblicati fra il 1877 e il 1897, agli otto articoli dal titolo comune *Symbolical reasoning* (*Ragionamento simbolico*), pubblicati su «Mind» a partire dal 1880 e fino al 1906; degno di menzione è anche il suo volume *Symbolic logic and its applications* (*Logica simbolica e sue applicazioni*) del 1906.

Da una concezione psicologistica espressa ad esempio dalla definizione di logica (del 1880) come «scienza generale del ragionamento considerato nel senso più astratto... scienza del ragionamento considerata in riflesso a quelle leggi e a quei principi generali del pensare che valgono quale che sia la natura del pensiero», McColl passa a una concezione nella quale la preminenza viene decisamente assegnata al momento segnico-calcolatorio, quando, nel 1897, afferma: «La logica simbolica (inclusa la matematica) può essere definita come la "scienza del ragionare mediante simboli rappresentativi"; questi simboli sono impiegati come sostituti per espressioni più lunghe di uso frequente», e afferma esplicitamente di voler mantenere logica e psicologia il più possibile separate (senza che ciò abbia a significare una totale indipendenza fra di esse) assumendosi in particolare lo scopo di riportare l'inferenza logica a una esecuzione meccanica.

Nel contesto di questa sua concezione «linguistica» McColl anticipa indubbiamente Frege nel ritenere che la logica possa e debba superare le «accidentalità» del linguaggio comune (ad esempio la distinzione soggetto-predicato) e globalmente sviluppa un sistema di logica enunciativa molto simile a quello che può oggi trovarsi in un testo introduttivo di logica. Intravvede la possibilità di una generalizzazione al momento predicativo (e non può evidentemente avanzare su questa strada non possedendo il concetto di funzione enunciativa, in quel periodo adombrato da Peirce ed esplicitamente sistemato da Frege) come pure sono chiari nei suoi scritti, e forse proprio in dipendenza dei suoi interessi «probabilistici», accenni che possono far pensare alla possibilità di logiche modali e libere (in cui non si assume che tutti i termini denotino

oggetti esistenti).[30] In particolare è proprio criticando l'implicazione materiale utilizzata da Russell e di cui parleremo più avanti che McColl reintroduce nel 1905 le distinzioni modali e dà inizio ad un dibattito che si riaprirà più tardi con Lewis sui rapporti tra inferenza condizionale e logica modale. Per McColl infatti il condizionale «se $\mathscr{A}$ allora $\mathscr{B}$» non si può rendere come fa Russell come «non $\mathscr{A}$ oppure $\mathscr{B}$» perché questo renderebbe validi legami tra gli enunciati che non sono logicamente necessari ma dipendono dalla specificazione delle circostanze in cui si realizzano. Ad esempio dati due enunciati qualsiasi $\mathscr{A}$ e $\mathscr{B}$, uno di essi implicherebbe l'altro, entrambi sarebbero implicati da ogni enunciato falso e implicherebbero ogni enunciato vero. È proprio per rendere questa indipendenza dalle circostanze che McColl introduce un operatore di necessità e interpreta «se $\mathscr{A}$ allora $\mathscr{B}$» come «è necessario che non $\mathscr{A}$ oppure $\mathscr{B}$». Il calcolo di McColl è formulato in termini «algebrici» (booleani) assumendo simboli quali A, B, C, ...; a, b, c, ...; α, β,γ, ...; x,y, z, ... come «variabili» per enunciati e quindi simboli «costanti»: «1» e «0» per indicare la verità, rispettivamente la falsità degli enunciati; «$\times$» o la semplice giustapposizione per il prodotto fra enunciati (congiunzione);[31] «$+$» per la «somma non esclusiva» (disgiunzione alternativa); «$'$» per la negazione; «:» per l'implicazione; «$=$» per l'equivalenza. Dà quindi 17 re-

[30] Ci sembra interessante e «istruttivo» riportare a questo punto la «valutazione» che, nella citata opera del 1884, Jevons dà del sistema di McColl. Afferma dunque Jevons: «Molto di recente Mr. Hugh McColl, B.A., ha pubblicato... numerosi lavori su un calcolo degli enunciati equivalenti... Il suo calcolo differisce in molti punti tanto da quello di Boole quanto da quello descritto in questo volume come logica equazionale. Mr. McColl respinge le equazioni a favore delle *implicazioni*; così la mia $A = AB$ diventa per lui $A : B$ o A implica B. Anche i suoi termini letterali differiscono in significato dai miei, poiché le sue lettere denotano proposizioni, non cose. Così $A : B$ asserisce che l'enunciato A implica l'enunciato B, o che ogniqualvolta A è vero anche B è vero. È difficile credere che vi sia alcun vantaggio in queste innovazioni... Mi sembra che i suoi propositi tendano verso un ricacciare indietro la logica formale nello stato di confusione in cui si trovava prima di Boole (*into its ante-boolian confusion*)».

[31] Riportiamo a mo' d'esempio una definizione e una regola dal *Calcolo di enunciati equivalenti*... del 1877. «Definizione 2. Il simbolo $A \times B \times C$ o ABC denota un enunciato composto del quale gli enunciati A, B, C sono detti i *fattori*. L'equazione $ABC = 1$ asserisce che *tutti e tre gli enunciati sono veri*; l'equazione $ABC = 0$ asserisce che almeno *uno* dei tre è falso...». «Regola 11. Se $A : B$ allora $B' : A'$. Così le implicazioni $A : B$ e $B' : A'$ sono equivalenti, ognuna di esse seguendo come necessaria conseguenza dall'altra. Questo è il principio logico di "contrapposizione"».

gole (fra cui ad esempio: distributività di «$\times$» rispetto a «+», leggi di De Morgan, legge di contrapposizione, ecc.) mediante le quali operare col sistema simbolico così istituito. McColl ottiene nel suo «calcolo» numerose «tesi» della logica proposizionale, la quale viene qui a essere per la prima volta «emancipata» dal calcolo delle classi, anticipando quanto lo stesso Russell farà nel 1906 nel suo articolo *The Theory of Implication* (*La teoria dell'implicazione*).

6. I CONTRIBUTI DI PEIRCE E DI SCHRÖDER

I motivi per cui abbiamo dedicato un paragrafo a parte all'esposizione dei contributi dell'americano Peirce e del tedesco Schröder sono duplici: da un lato, a giustificare perché li abbiamo «separati», occorre considerare che i loro apporti sono di gran lunga i più incisivi e significanti per lo sviluppo dell'algebra della logica, vuoi per l'originalità, l'acutezza e la varietà (Peirce) vuoi per la sistematica organicità (Schröder); in certo senso essi costituiscono un punto d'arrivo, una *summa* completa di quello che è lo stato della logica algebrica ottocentesca. D'altro lato, a giustificare perché li abbiamo «uniti», diciamo che ci sembra interessante far risaltare due tipi di approccio assolutamente diversi a una stessa problematica.

Peirce e Schröder infatti, pur lavorando indipendentemente, hanno stretti rapporti scientifici fra loro: Peirce legge e recensisce le opere di Schröder, quest'ultimo traduce operativamente i vari e notevoli spunti che le riflessioni di Peirce facevano emergere. Ma tanto è sistematico, pignolo, «tedesco» Schröder, quanto è disordinato, geniale e frammentario Peirce. Questi può considerarsi il vero e proprio *trait-d'union* fra la logica algebrica e quella logicista, e ciò indubbiamente in conseguenza di una impostazione aperta e autocritica di tutta la sua speculazione; quello invece è convinto della verità e unicità della logica booleana, che riesce sì a esporre con un alto grado di sistematicità e «completezza», ma entro i limiti appunto di una quasi dogmatica convinzione di aver a che fare con *il* metodo di ragionamento assoluto e definitivo. Cominciamo quindi con l'esporre i contributi di Peirce, per poi illustrare brevemente la sistemazione schröderiana.

Charles Sanders Peirce[32] si dichiara profeticamente convinto che la logica farà progressi enormi rispetto allo stato in cui egli la lascia ed è altresì convinto di contribuire in modo sostanziale a tale processo di sviluppo. In effetti i contributi di Peirce alla logica formale sono numerosi e assai profondi e possono essere riassunti in tre rubriche, inquadrate in una visione unitaria di tutta la problematica logica, che è caratteristica di Peirce e che prelude a una generale sistemazione assiomatica dei vari calcoli: 1) approfondimento e sistemazione del calcolo booleano, nel più ampio contesto di una chiarificazione del rapporto fra logica e matematica; 2) interpretazione proposizionale dello stesso, con notevoli «anticipazioni» di risultati ottenuti indipendentemente da lui, da Frege e da Sheffer e Wittgenstein; 3) impostazione e sviluppo decisivo della logica delle relazioni, anche qui con notevoli anticipazioni dell'opera di Russell e Frege e quindi, in generale, della sistemazione logicista della logica.

Prenderemo separatamente in esame questi tre campi nell'ordine sopra esposto; per comodità del lettore, invece di riferirci semplicemente ai *Collected papers*, indichiamo qui gli scritti di Peirce di cui ci serviremo nella nostra esposizione, e che elencheremo in ordine cronologico: *On a improvement in Boole's e calculus of logic* (*Su un perfezionamento del calcolo della logica di Boole*, 1867, che citeremo nel seguito con P1867[1]); *Upon the logic of mathematics* (*Sulla logica della matematica*, 1867, citato con P1867[2]); *Description of a notation for the*

[32] Figlio del valente matematico Benjamin (1809-80), il filosofo e logico americano Charles Sanders Peirce nacque a Cambridge (Mass.) nel 1839 e morì a Milford (Penn.) nel 1914. Ad Harvard, dove insegnava il padre, divenne *master of arts* nel 1862 e *bachelor of science* nel 1863. Per circa trent'anni fu impegnato in ricerche statistiche per conto della U.S. Coast and Geodetic Survey, a partire dal 1861, dirigendo nel contempo il Dipartimento americano dei pesi e delle misure. Non si dedicò formalmente all'insegnamento, ma fu *lecturer* di logica a Harvard nel 1864-65 e nel 1869-70; quindi alla Johns Hopkins University di Boston fra il 1879 e il 1884 e dal 1891 in poi. Scrittore prolifico e asistematico, ebbe interessi diversi, dalla storia della scienza alla metafisica, dalla logica alla matematica e all'astronomia, alla chimica. Molti suoi saggi e articoli comparvero nelle più diffuse riviste filosofiche e scientifiche dell'epoca, ma la maggior parte dei suoi scritti restarono inediti e furono raccolti postumi: una prima ristretta scelta nel volumetto *Chance, love and logic* (*Caso, amore e logica*) del 1923, quindi nei *Collected papers of Charles Sanders Peirce* (*Tutti gli scritti di C.S. Peirce*) in otto volumi: dal primo al sesto a cura di Charles Hartshorne e Paul Weiss fra il 1931 e il 1935; il settimo e l'ottavo a cura di Arthur W. Burks nel 1958. I volumi sono così intitolati: I *Principles of philosophy* (*Principi di filosofia*); II *Elements of logic* (*Elementi di logica*); III *Exact*

logic of relatives, resulting from an amplification of the conceptions of Boole's calculus of logic (*Descrizione di una notazione per la logica dei relativi, risultante da un ampliamento delle concezioni del calcolo logico di Boole*, 1870, P1870); un articolo pubblicato postumo nel 1933, ma risalente circa al 1880 (e che indicheremo con P1880[1]), *A boolian algebra with one constant* (*Un'algebra booleana con una costante*); *On the algebra of logic* (*Sull'algebra della logica*, 1880, P1880[2]); *The logic of relatives* (*La logica dei relativi*, 1883, P1883); *On the algebra of logic: a contribution to the philosophy of notation* (*Sull'algebra della logica: un contributo alla filosofia della notazione*, 1885, P1885); *The regenerated logic* (*La logica rigenerata*, 1896, P1896); *The logic of relatives* (*La logica dei relativi*, 1897, P1897); un altro inedito *The simplest mathematics* (*La matematica più semplice*, 1902, P1902).

1) Come si ricorderà nel suo sistema Boole «esprimeva» quattro operazioni logiche: la somma esclusiva « + », il prodotto « × » (in generale indicato per semplice giustapposizione) e le loro inverse « − » e « / », definite in modo che se $x + b = a$ allora $x = a - b$ e se $xb = a$ allora $x = a/b$. Queste operazioni si comportano diversamente rispetto alle analoghe operazioni aritmetiche. Convenendo di indicare con $[n]x$ il numero degli elementi della classe x, si ha

$$[n]a + [n]b = [n](a + b)$$

logic (*Logica esatta*); IV *The simplest mathematics* (*La matematica più semplice*); V *Pragmatism and pragmaticism* (*Pragmatismo e pragmaticismo*); VI *Scientific metaphysics* (*Metafisica scientifica*); VII *Science and philosophy* (*Scienza e filosofia*); VIII *Reviews. Correspondance* (*Recensioni. Corrispondenza*). L'asistematicità tipica del Peirce e i criteri di distribuzione dei volumi rendono la lettura della sua opera alquanto difficoltosa. Oggi è in corso una nuova edizione (di cui sono comparsi i primi cinque volumi) con criteri più corretti e che si annuncia come estremamente interessante. Nel 1893 Peirce aveva portato a termine, pur lasciandola inedita, una grande opera *The grand logic* (*La grande logica*) o *How to reason: a critic of arguments* (*Come ragionare: una critica dell'argomentazione*) e verso il 1902 aveva steso tre capitoli e l'inizio di un quarto di una seconda opera organica la *Minute logic* (*Logica minuta*) che rimase però incompiuta e inedita. Questa incapacità di portare a termine i numerosi progetti intrapresi, accanto all'originalità e genialità universalmente riconosciutegli, è uno dei motivi per cui spesso Peirce viene avvicinato a Leibniz. È stato peraltro detto che se «Peirce avesse avuto il piacevole e discorsivo stile di De Morgan o il metodo dettagliato e accurato di Schröder, la sua opera sulla logica simbolica avrebbe occupato numerosi volumi». Peirce è universalmente considerato il fondatore del pragmatismo americano.

e ciò in dipendenza delle condizioni di «esclusività» poste da Boole per l'operazione di somma logica. Si controlla facilmente che la stessa relazione vale per l'inversa « — », e si può concludere che « + » e « — » differiscono dalle omonime aritmetiche per il solo fatto che ammettono argomenti che invece di essere soltanto numeri sono in generale classi o comunque non necessariamente «quantità». Comportamento del tutto diverso hanno invece l'operazione di prodotto logico e la sua inversa. In questo caso, per esempio, la relazione

$$[n]a \times [n]b = [n](a \times b)$$

vale solo se *a* e *b* assumono i valori 0 e/o 1 e lo stesso vale per l'analoga relazione scritta per «/»; sicché queste operazioni, per così dire, non hanno «controparte» aritmetica.

Ora Peirce, consapevole tra l'altro della critica di Jevons contro l'esclusività della somma logica booleana, pensa di ovviare a quello che anch'egli ritiene un difetto del sistema di Boole, con un'impostazione del tutto originale, «completando» nel sistema stesso la «parte» logica e quella aritmetica. Queste variazioni sono motivate dalla ricerca di un chiarimento del rapporto fra il dominio delle operazioni logiche e matematiche, in definitiva dall'esigenza di indagare, in generale, il rapporto logica-matematica: «Queste considerazioni,» dirà infatti Peirce in P1867[2], « ... porranno, spero, le relazioni fra logica e matematica in una luce in qualche modo più chiara che nel passato.» Peirce basa così il suo sistema su due tipi di operazioni: quelle logiche e quelle aritmetiche. Le quattro operazioni logiche o non aritmetiche, sono:

a) la somma, che egli indica con « +,» in modo tale che «*a* +, *b* denoti tutti gli individui contenuti sotto *a* e *b* assieme. L'operazione qui eseguita differirà in due rispetti dall'addizione aritmetica: 1) essa si riferisce a identità, non a uguaglianze; 2) ciò che è comune ad *a* e a *b* non viene considerato due volte come succederebbe in aritmetica»;

b) il prodotto, che egli indica con *a* ×, *b* o semplicemente con *a*, *b* e che coincide col prodotto booleano;

c-d) le operazioni logiche inverse delle precedenti che noi comunque non considereremo per motivi che appariranno subito chiari.

Le quattro operazioni «aritmetiche» sono invece:

a') la somma « + » che coincide con la somma booleana ossia è esclusiva, e la sua conversa « — »;

b') il prodotto $a \times b$ e il suo inverso $a \div b$.

Questo modo di affrontare il problema del disancoramento dell'algebra di Boole da un'eccessiva aderenza alle analogie matematiche viene elaborato in P1867[1]; tuttavia già in P1867[2] Peirce conserva solo le operazioni dirette di somma e prodotto: «Poiché le operazioni inverse non hanno alcun interesse logico peculiare esse vengono qui trascurate.» In effetti in P1870 egli mantiene ancora le varie operazioni ad eccezione di $\div$ ma in articoli posteriori tutto il calcolo viene impostato sulle sole operazioni di somma e prodotto (logici, naturalmente). È chiaro che in questo modo Peirce può assumere le due leggi «critiche» $a +, a = a$ e $a, a = a$ ottenendo così in particolare la semplificazione e la simmetria che, come sappiamo, Jevons aveva richieste per primo.

Un ulteriore passo verso quella che possiamo chiamare la «disaritmetizzazione» del calcolo di Boole, viene compiuto da Peirce nel citato P1870 con l'introduzione della relazione di inclusione, che egli indica col simbolo $-\!\!<$ e definisce come segue: «*inclusione* o *essere tanto piccolo quanto* è una relazione transitiva. Vale la conseguenza che se $x -\!\!< y$ e $y -\!\!< z$ allora $x -\!\!< z$». Osservando che l'unica relazione booleana, la « = », può essere espressa in termini dell'inclusione (ponendo per definizione che $x = y$ equivalga a $x -\!\!< y$ e $y -\!\!< x$), si può concludere che già nel 1870 Peirce disponeva di una sistemazione che possiamo ritenere definitiva (rispetto ovviamente alle sistemazioni oggi in uso) del calcolo delle classi. A ragione quindi Lewis afferma che «a parte il trattamento dei termini relativi e l'uso di operazioni "aritmetiche", questa monografia [la P1870] dà le leggi della logica delle classi sostanzialmente in modo identico all'attuale algebra della logica».

Un ulteriore progresso in questo senso compie del resto Peirce in P1880[2], quando la relazione di inclusione viene sistematicamente posta come fondamento del calcolo, definendo le due operazioni di somma e prodotto logico in termini di inclusione, osservando che «se $a -\!\!< x$ e $b -\!\!< x$ allora $a +, b -\!\!< x$ e inversamente, se $a +, b -\!\!< x$, allora $a -\!\!< x$ e $b -\!\!< x$; se $x -\!\!< a$ e $x -\!\!< b$ allora $x -\!\!< a,b$ e inversamente se $x -\!\!< a,b$, allora $x -\!\!< a$ e $x -\!\!< b$».

Privilegiando la relazione $-\!\!<$ non solo Peirce anticipa quella che sarà poi la trattazione reticolare delle algebre di Boole, ma,

come anche McColl, assumendo l'implicazione come relazione fondamentale al posto dell'uguaglianza, segna un deciso avvicinamento fra la tradizione algebrica e l'approccio logicista di Frege.

2) Si può comprendere così come l'introduzione della relazione $-\!\!<$ costituisse un passo decisivo rispetto alla sistemazione booleana. Peirce prende certamente lo spunto per la sua ricerca dalle difficoltà cui andava incontro l'interpretazione del calcolo booleano in termini di proposizioni «secondarie» e in P1880[2] comincia appunto con l'osservare che la relazione $A -\!\!< B$ – che ora va intesa come «A implica B» o «se A è vero allora B è vero» – può sussistere tanto fra proposizioni primarie («categoriche») quanto fra proposizioni secondarie («ipotetiche»). La codificazione della «logica non relativa» o «logica dei termini assoluti» come calcolo proposizionale avviene in P1885, ove Peirce assume come fondamentale per la sua sistemazione la relazione $-\!\!<$ intesa come *implicazione materiale* (o filoniana); Peirce argomenta a lungo, adducendo anche motivazioni di tipo storico, per mostrare come questo tipo di implicazione sia più adeguata ai suoi scopi di quella diodorea (o implicazione stretta) e nel contempo mette in luce con critica consapevolezza i limiti intrinseci dell'implicazione materiale per quanto riguarda la traduzione formale delle varie possibili accezioni dell'ordinario termine «implica».

In P1885 il calcolo viene impostato servendosi del *modus ponens* e del principio di sostituzione quali regole di inferenza, ossia dando una sistemazione sostanzialmente analoga a quella che Frege aveva già dato, come vedremo, nel 1879. Il fatto di essere stato «preceduto» da Frege non toglie certo nulla al valore delle intuizioni e delle elaborazioni di Peirce, il quale da una parte (almeno per quanto siamo riusciti ad accertare) era all'oscuro dell'opera del logico tedesco, e dall'altra era giunto alla sua sistemazione a partire da una problematica algebrica, sostanzialmente diversa cioè da quella di Frege.

D'altronde altri esempi significativi di come Peirce fosse giunto a conclusioni e risultati che sembrerebbero (e che da molti sono ancor oggi ritenuti) caratteristici di una impostazione logicista, non mancano certamente. Basti pensare ad esempio al breve saggio (cinque pagine) P1880[1] motivato come segue: «L'apparato del calcolo booleano consiste dei segni $=$, $>$ (non usato da Boole, ma necessario per esprimere proposizioni particolari), $+$, $\times$, $-$, 1, 0.

Al posto di questi sette segni io propongo di impiegarne uno solo», che è una chiara anticipazione della scoperta, che Sheffer farà nel 1913, del fatto che un solo connettivo binario, detto appunto *stroke* (o «sbarra») di Sheffer, basta a esprimere ogni funzione di verità. O ancora, sarebbe sufficiente considerare, in P1902, la descrizione che Peirce dà del concetto di tavola di verità (anticipazione questa di Schröder e poi di Wittgenstein) che viene anzi addirittura impiegato, già in P1885, come procedura di decisione della «tautologicità» di una formula proposizionale con un metodo che noi oggi chiameremmo del controesempio e che anticipa le tavole di Beth. Esso consiste, per dimostrare che una formula è tautologica, nel supporre che essa possa risultare falsa per un'assegnazione di valori di verità, e nel mostrare quindi che tale ipotesi porta a contraddizione. Mette conto di riportare il breve passo di Peirce: «Una proposizione della forma $x -\!\!< y$ è vera se $x = \text{f}$ o $y = \text{v}$. Una proposizione della forma $(\overline{x -\!\!< y})$ è vera se $x = \text{v}$ e $y = \text{f}$. Di conseguenza, per trovare se una formula è necessariamente vera [è una tautologia] si sostituiscano v e f alle lettere e si veda se può essere supposta falsa mediante ognuno di tali assegnamenti di valore. Prendiamo ad esempio la formula

$$(x -\!\!< y) -\!\!< \{(y -\!\!< z) -\!\!< (x -\!\!< z)\}.$$

Per renderla falsa dobbiamo prendere

$$(x -\!\!< y) = \text{v}; \{(y -\!\!< z) -\!\!< (x -\!\!< z)\} = \text{f}.$$

L'ultima dà

$$(y -\!\!< z) = \text{v e } (x -\!\!< z) = \text{f}, x = \text{v}, z = \text{f}.$$

Sostituendo questi valori in

$$(x -\!\!< y) = \text{v e } (y -\!\!< z) = \text{v},$$

abbiamo

$$(\text{v} -\!\!< y) = \text{v e } (y -\!\!< \text{f}) = \text{v}$$

che non possono essere contemporaneamente soddisfatte».

3) I contributi più rilevanti di Peirce riguardano però la logica delle relazioni (o dei «relativi» in terminologia peirciana). Peirce prende certamente le mosse dal quarto degli articoli di De Morgan sul sillogismo, pubblicato come si ricorderà nel 1860, trovandolo insufficiente sotto molti riguardi e nota d'altra parte che sarebbe estremamente interessante tentare di estendere l'algebra di Boole – che «ha sì singolare bellezza» – «all'intero campo della logica formale, invece di lasciarla ristretta a quella parte elementare e meno utile della materia, la logica dei termini assoluti, che

quando egli [Boole] scriveva era l'unica logica formale conosciuta». Così si legge in P1870, progettato appunto per presentare operativamente tale estensione. Altri lavori di Peirce specificamente dedicati alla logica delle relazioni sono P1880[2], P1883, P1885, P1896, P1897; noi ci riferiremo essenzialmente a P1883 (e in parte a P1885) che si può ritenere contenga la versione definitiva, anche da un punto di vista simbolico, del sistema logico-relazionale peirciano.[33] Il discorso è in gran parte condotto esemplificandolo per relazioni binarie (o «duali» come dice Peirce) ma il simbolismo adottato definitivamente da Peirce (sostanzialmente più perspicuo e maneggevole di quello impiegato in lavori precedenti) porta in modo naturale a generalizzare a «relativi tripli» o con un numero finito qualunque di posti. Comunque nel seguito, se non espressamente indicato, con «relativo» (o, in termini moderni, «relazione») intenderemo «relativo binario».

Peirce definisce un «termine relativo duale» come «un nome significante una coppia d'oggetti» vale a dire identifica un «relativo individuale» con la *coppia ordinata* costituita dal «relato» e dal «correlato»; così ad esempio considerando la relazione «... è benefattore di - - -» i puntini segnano il posto del relato e le lineette quelle del correlato (nel seguito diremo semplicemente la «relazione di benefattore»). Supponiamo ora, con Peirce, che *A*, *B*, *C*, *D*, ... siano gli individui dell'universo del discorso. Consideriamo

[33] Questo volume è importante, oltre che per il lavoro di Peirce cui ci riferiremo nel seguito del testo, anche come testimonianza dell'attività della scuola logica che Peirce aveva radunato attorno a sé alla Johns Hopkins University in quel periodo. Appare col titolo *Studies in logic. By members of the Johns Hopkins University* (*Studi logici, dei membri dell'università Johns Hopkins*) e converrà dare una descrizione succinta del suo contenuto. Esso contiene sei saggi, l'ultimo dei quali, dello stesso Peirce, è corredato da due «note» o appendici. Il primo saggio è di Allan Marquand e porta il titolo *The logic of epicureans* (*La logica degli epicurei*); in esso l'autore sostiene l'esistenza presso gli epicurei di «una logica che superava in valore quella dei loro rivali stoici». Si tratta di una logica induttiva, che Marquand avvicina a quella di J.S. Mill, e della quale dà i «canoni». Il secondo articolo, dello stesso autore, porta il titolo *A machine for producing syllogistic variations* (*Una macchina per produrre variazioni sillogistiche*) e, come lo stesso titolo fa intuire, in esso Marquand presenta un progetto per la costruzione di una semplice macchina basata sul sistema binario di numerazione, che permette di ottenere gli «otto modi di sillogismo universale di De Morgan» mediante rotazioni di alcuni cilindri, che equivalgono a compiere trasformazioni dirette sulle proposizioni categoriche, come ad esempio la contrapposizione. In una brevissima «nota» l'autore annuncia quindi di aver esteso il suo metodo a proposizioni con quattro e otto termini. Il quarto

tutte le coppie ordinate (relativi individuali) di tali individui, disponendole per comodità in una matrice come la seguente

$A : A$	$A : B$	$A : C$	$A : D$	...
$B : A$	$B : B$	$B : C$	$B : D$	...
$C : A$	$C : B$	$C : C$	$C : D$	...
$D : A$	$D : B$	$D : C$	$D : D$	...
...	...	...	...	...

Possiamo allora definire con Peirce un «relativo generale» come «l'aggregato logico di un numero qualsiasi di tali relativi individuali». Se per esempio l è la relazione «... è amante di ---» (*lover* in inglese), si potrà scrivere

$$l = \Sigma_i \Sigma_j \, (l)_{ij} \, (I : J)$$

«dove $(l)_{ij}$ è un coefficiente numerico il cui valore è 1 nel caso in cui I sia amante di J, 0 nel caso contrario, e dove le somme vanno estese a tutti gli individui del dominio». Si noti che questa definizione non comporta, come si potrebbe essere portati a credere, l'identificazione da parte di Peirce di una relazione con la *classe* delle coppie ordinate degli individui del dominio che stanno in quella relazione; un relativo generale, secondo la definizione peirciana, sta piuttosto per un relativo individuale (ossia per *una* coppia) non

saggio è estremamente significativo: l'autrice, Christine Ladd, sotto il titolo *On the algebra of logic* (*Sull'algebra della logica*), propone in realtà una trattazione esaustiva del sistema booleano che grazie all'impiego di una nuova notazione presenta in effetti vari vantaggi di simmetria, semplicità e maneggevolezza, ma che tuttavia non è stata accettata negli sviluppi successivi della logica. Ancora alla logica deduttiva è dedicato il quinto saggio del volume, dal titolo *On a new algebra of logic* (*Su una nuova algebra della logica*). Ne è autore O.H. Mitchell cui Peirce darà esplicitamente il merito di aver indicato come «si debba fare» per introdurre i quantificatori nel discorso logico. Qui ci limiteremo a dire, con lo stesso Peirce, che in questo saggio Mitchell «presenta nuovi sviluppi dell'algebra logica di Boole». Segue quindi l'articolo di B.I. Gilman *Operations on relative numbers with applications to the theory of probabilities* (*Operazioni sui numeri relativi con applicazioni alla teoria delle probabilità*), nel quale l'autore sviluppa quelle regole per la combinazione di numeri relativi (relativi nel senso che la loro definizione è strettamente collegata con la teoria delle relazioni) di cui i principi generali della teoria delle probabilità sono casi particolari. Il contributo di Peirce è infine triplice: all'articolo *A theory of probable inference* (*Una teoria dell'inferenza probabile*) seguono la nota A, *On a limited universe of marks (Su un universo limitato di segni)*, e la nota B, *The logic of relatives* (*La logica dei relativi*), che abbiamo già citato nel testo come P1883.

specificato, per un *qualunque* relativo individuale della classe. È anche opportuno notare subito che neppure Peirce riesce a evitare l'ambiguità che già in De Morgan abbiano riscontrato, e cioè quella di usare uno stesso simbolo per rappresentare tanto la relazione quanto gli individui che stanno nella relazione stessa. Si noti ancora che a meno di minori precisazioni, si può intendere un coefficiente come l_{ij} come una funzione proposizionale (altro concetto «logicista» chiaramente anticipato da Peirce) ove gli indici svolgono la funzione di variabili; l_{ij} non è cioè una proposizione, ma afferma che l'individuo genericamente indicato con i è amante dell'individuo genericamente indicato con j. Diventa una proposizione (vera o falsa) quando al posto di i e j si mettano i nomi di due individui (non necessariamente diversi) di un dato dominio.

Peirce estende quindi ai termini relativi le ordinarie operazioni di somma e prodotto logici note dall'algebra dei termini assoluti. Le definizioni sono ovviamente, per la somma « + »

$$(l + b)_{ij} = (l)_{ij} + (b)_{ij}$$

e per il prodotto «,»

$$(l, b)_{ij} = (l)_{ij}(b)_{ij}$$

dove nella prima, tenendo conto del significato dei simboli, il segno « + » a destra del segno di uguaglianza va interpretato come quella operazione così definita:

$$0 + 0 = 0, 0 + 1 = 1 + 0 = 1 + 1 = 1.$$

Parlando in termini moderni, interpretando al solito l come la relazione di amante e b come la relazione di benefattore, le operazioni precedenti sono le ordinarie operazioni booleane: la prima sta a significare la relazione ottenuta dalla disgiunzione delle due date, ossia la relazione «amante *oppure* benefattore» (con «oppure» inteso in senso inclusivo) mentre la seconda indica la relazione ottenuta dalla congiunzione delle due date, ossia «amante *e* benefattore». Viene anche estesa ai termini relativi l'operazione detta di «negazione» in termini peirciani, e che noi diciamo (insiemisticamente) di complemento: essa viene indicata soprassegnando il segno della relazione data e include ogni coppia esclusa per così dire dalla relazione «positiva». Così, se l è al solito la relazione di amante, $\bar{l}$ è la relazione di non-amante.

Vengono poi introdotte operazioni caratteristiche dei termini relativi, non applicabili cioè a termini assoluti: l'operazione unaria di *conversione* applicata a una relazione l ne dà il *converso*, ossia quella relazione, indicata con $\breve{l}$, che si ottiene invertendo l'ordine delle coppie nella relazione data; se al solito l sta per la relazione di amante di, il suo converso sarà la relazione di amato da. Quindi vengono definite, in termini di operazioni booleane, due operazioni relative binarie: la prima detta di *prodotto relativo* (che abbiamo già incontrato in De Morgan) viene indicata per semplice giustapposizione ed è così definita

$$(lb)_{ij} = \Sigma_x (l)_{ix} (b)_{xj};$$

la seconda, detta *somma relativa*, viene indicata con † ed è definita con l'equazione

$$(l \dagger b)_{ij} = \Pi_x (l)_{ix} + (b)_{xj}$$

col significato di « + » sopra definito. La seconda di queste operazioni viene interpretata da Peirce come «l di tutti i non-b di...» ed è di scarsa se non nulla utilità; probabilmente è stata introdotta da Peirce per mantenere la validità delle leggi di De Morgan anche per queste operazioni relative; la prima invece è un'operazione estremamente interessante e di grande importanza, e va interpretata come «l di un b di...» ossia, più chiaramente, sussiste fra due individui x e y se (e solo se) esiste un altro individuo z tale che «l_{xz} e b_{zy}». Si pensi ad esempio alla relazione «x è zio paterno di y»: essa sussiste fra x e y se e solo se esiste uno z tale che x è fratello di z e z è padre di y, vale a dire tale relazione è prodotto relativo delle relazioni di fratello e di padre.

Le operazioni relative binarie godono della proprietà associativa ma in generale non di quella commutativa. Esse hanno inoltre un comportamento fra loro asimmetrico come messo in evidenza dalle seguenti leggi, che Peirce dimostra fra molte altre:

$$\begin{aligned} & ls + bs \prec (l + b)s \\ & l, b \dagger s \prec (l \dagger s), (b \dagger s) \\ & (l + b)\, s \prec ls + b \\ & (l \dagger s), (b \dagger s) \prec l, b \dagger s. \end{aligned}$$

Vengono quindi date le leggi che regolano il mutuo comportamento delle somme e dei prodotti relativi e assoluti rispetto alla negazione e alla conversione e si definiscono le relazioni riflessive (*selfrelatives*) e non riflessive (*aliorelatives*). Peirce introduce quindi quattro relazioni particolari: la relazione universale, simbolizzata con ∞, resa in linguaggio ordinario con «coesistente con»: ∞ accoppia ogni elemento del dominio con se stesso e con ogni altro; la negazione di tale relazione, indicata con 0; la relazione «identico con», indicata con 1, che accoppia ogni termine con se stesso; la negazione di tale relazione indicata con n, che accoppia ogni termine con ogni altro termine da esso distinto. Anche per queste relazioni Peirce stabilisce diverse leggi, come ad esempio $0 \prec x$, $x \prec \infty$ qualunque sia x, $x0 = 0$, $x \dagger \infty = \infty$, e molte altre, che ci asterremo dal riportare.

È chiaro che la maggior quantità di operazioni che intervengono nel calcolo, congiunta con l'assenza di simmetria nel loro comportamento rendono il calcolo stesso assai complesso; Peirce afferma infatti che «la logica dei relativi è altamente multiforme; essa è caratterizzata da innumerevoli inferenze immediate e da varie conclusioni distinte dallo stesso insieme di premesse... L'effetto di queste peculiarità è che quest'algebra non può essere soggetta a rigide e solide regole come quelle del calcolo booleano; e tutto ciò che si può fare in questa sede è dare un'idea generale di come lavorare con essa». In proposito Lewis osserva che «... come oggi sappiamo, il valore principale di ogni calcolo delle relazioni non consiste nelle eliminazioni o soluzioni di tipo algebrico, ma in deduzioni da farsi direttamente dalle sue formule. Gli artifici di Peirce per la soluzione sono quindi molto meno importanti della fondazione teoretica sulla quale è costruito il suo calcolo dei relativi. È questa che si è dimostrata utile in ricerche successive ed è divenuta la base di notevoli apporti allo sviluppo logicista».

Un aspetto per noi estremamente interessante di questo è costituito dagli accenni già ampiamente contenuti in P1883, e poi sviluppati abbastanza organicamente in P1885, di una vera e propria logica della quantificazione. In P1883, gli operatori Σ e Π avevano significati, anche se non esclusivamente, almeno sostanzialmente algebrici, che tuttavia permettevano a Peirce affermazioni come la seguente: «Quando le operazioni relative e quelle non relative [booleane] occorrono assieme, le regole del calcolo diventa-

no molto complicate. In questi casi, come pure nei casi in cui intervengano relazioni *plurali* (che sussistono fra tre o più oggetti) è spesso vantaggioso ricorrere ai coefficienti numerici... *Ogni proposizione è equivalente a dire che qualche complesso di somme e prodotti di tali coefficienti numerici è maggiore di zero.* Così

$$\Sigma_i \Sigma_j l_{ij} > 0$$

significa che qualcosa è amante di qualcosa; e

$$\Pi_i \Sigma_j l_{ij} > 0$$

significa che ogni cosa è amante di qualche cosa. Possiamo, naturalmente, *omettere nella notazione delle disuguaglianze, il* $>$ 0 *con le quali tutte esse terminano* [corsivo nostro]». Risulta chiaro l'impiego di tali operatori come quantificatori «qualche» e «tutti» e già in questo saggio viene *in nuce* usata la distinzione fra parte «booleana» (noi diremmo oggi proposizionale o «matrice») e parte «relativa» (noi diremmo di quantificazione o «prefisso») di una espressione, e vengono dati alcuni esempi di trasformazione del «prefisso», e quindi di deduzioni, che possono automaticamente interpretarsi nell'ambito dei predicati del primo ordine. Si ricordi del resto che i coefficienti relativi introdotti da Peirce altro non sono, sostanzialmente, che le funzioni proposizionali di Frege e di Russell.

Nel saggio del 1885 comunque, una teoria della quantificazione viene impostata con tutta chiarezza: viene tra l'altro considerata esplicitamente la possibilità di porre una espressione in quella che noi oggi chiamiamo forma prenessa (nella quale cioè tutti i quantificatori precedono tutti i connettivi) e vengono anche eseguite forme particolari di tale trasformazione. Afferma infatti Peirce in P1885: «Veniamo ora alla distinzione di *qualche* e *tutti*, distinzione che è precisamente accoppiata con quella fra verità e falsità. Tutti i tentativi per introdurre questa distinzione nell'algebra booleana sono stati più o meno un fallimento completo, finché il signor Mitchell [in un suo saggio contenuto nello stesso volume nel quale si trova P1883; cfr. nota 33 a pag. 178] non ha mostrato come va effettuata. Il suo metodo consiste in realtà nel rendere l'intera espressione della proposizione consistente di due parti, una pura espressione booleana che si riferisce a un individuo e una parte

quantificante che dice di che individuo si tratti. Così, se k significa "egli è un re" e h "egli è felice", la booleana $(\bar{k} + h)$ significa che l'individuo di cui si parla o non è un re o è felice. Ora, applicando la quantificazione, possiamo scrivere Ogni $(\bar{k} + h)$ per significare che ciò è vero di ogni individuo dell'universo (limitato), oppure Qualche $(\bar{k} + h)$ per significare che esiste un individuo che o non è re o è felice.» Introdotti quindi i simboli Σ per «qualche» e Π per «ogni», si ha che «$\Sigma_i\, x_i$ significa che x è vero di qualcuno degli individui denotati da i, ossia $\Sigma_i\, x_i = x_i + x_j + x_k +$ ecc. Allo stesso modo, $\Pi_i\, x_i = x_i\, x_j\, x_k$ ecc. ... Va notato che Σ e Π sono soltanto *simili* a una somma e a un prodotto; non sono a rigore di quella natura perché gli individui dell'universo possono essere non numerabili».

Per concludere il nostro discorso su Peirce vogliamo toccare ancora due punti, il primo dei quali riguarda la polemica da lui condotta contro la tesi logicista sulla riducibilità della matematica alla logica. Abbiamo detto che tale tesi risale in modo esplicito e articolato a Frege, ma Peirce ne ha conoscenza attraverso l'opera *Was sind und was sollen die Zahlen?* (*Essenza e significato dei numeri*, 1888) di Dedekind, di cui avremo occasione di parlare più avanti. Ciò non deve stupire: si può senz'altro anticipare, incidentalmente, che i lavori e le idee di Frege rimasero sostanzialmente ignoti ai suoi contemporanei e, se noti, vennero assolutamente sottovalutati e del tutto trascurati. La cosa è imputabile a diversi fattori: da una parte la novità delle idee stesse, il tipo di presentazione fregeano, almeno in un primo momento alquanto ostico (con scarse se non nulle concessioni «alla platea») e al di fuori di una trattazione tradizionalmente matematica, lo spirito polemico di Frege, il suo carattere non certo accomodante e «socievole»; d'altra parte però non ultima delle ragioni di questo stato di cose è proprio la «chiusura» di logici pur rilevanti come Schröder, la loro incapacità a cogliere la novità e la fecondità di quelle idee per il semplice motivo che la logica cui Frege intendeva appunto ridurre la matematica non era *la* logica ossia quella booleana. Come vedremo in seguito, nel 1891 Frege aveva già ampiamente esposto, e in varie presentazioni, le proprie idee (basterebbe limitarsi ai lavori del 1879 e del 1884) e si apprestava a presentare operativamente il suo programma nel primo volume della sua opera maggiore i *Grundgesetze der Arithmetik* (*Principi dell'aritmetica*), che pubblicherà

nel 1893; orbene, ancora nel 1891, nel secondo volume delle *Vorlesungen*, Schröder scriveva: «L'aritmetica è ora completamente liberata dall'intuizione o dall'assunzione di grandezze confrontabili e misurabili ed è posta su fondamenti puramente logici; è soprattutto grazie ai lavori di H. Grassmann, G. Cantor, Weierstrass e Dedekind che essa è stata sviluppata come un ramo della logica»; la cosa si commenta da sola, tenendo anche conto del fatto che nei tre volumi della *summa* schröderiana Frege è sì e no nominato.

Ma a parte questa digressione, quali sono i motivi che portano un logico di interessi filosofici ben più ricchi come Peirce a non accettare e anzi a opporsi decisamente alla tesi logicista? Ciò non avviene per una dogmatica affermazione dell'unicità e della validità del modello matematico, ma perché matematica e logica vengono fondamentalmente intese da Peirce, in generale, come «ricerca strutturale sul simbolismo»; è quindi possibile e lecito effettuare una revisione critica delle teorie matematiche servendosi di tutto l'apparato simbolico e concettuale della logica, ma questa impresa non comporta assolutamente un momento fondante o comunque di riduzione della matematica alla logica, considerato appunto che nel senso preciso sopra esposto *anche* la logica è matematica. In altri termini, nel corso di questa revisione non si esce assolutamente dal campo stesso della matematica. In modo diverso si può dire che è impossibile una riduzione della matematica alla logica, dal momento che la matematica è la scienza che *trae* conclusioni necessarie (secondo una definizione del padre di Peirce) e non la scienza *del trarre* tali conclusioni. La stretta connessione che Peirce sostiene fra *logica utens* e *logica docens* non è assolutamente un'identificazione: la logica formale in quanto attività di ricerca operante è niente altro che la matematica che «elabora i suoi ragionamenti per mezzo di una *logica utens* che sviluppa di per sé»; la *logica docens*, viceversa è il momento teorico che riflette su questi ragionamenti e procede alla «classificazione» degli argomenti. Si noti infine – e ci limitiamo a enunciare la cosa senza adeguatamente giustificarla – che secondo Peirce una concezione logicista corre in più sensi il pericolo di riportare alla tematica delle «leggi del pensiero» tipica del resto della tradizione europea dell'algebra della logica.

Quest'ultima considerazione ci conduce direttamente al secondo punto che vogliamo trattare, e cioè il già accennato atteggia-

mento antipsicologistico di Peirce. In proposito Peirce osserva, tra l'altro, che «quest'opinione [psicologista] è sostenuta da uno dei logici odierni più acuti e alla moda, Cristoph Sigwart... e Schröder, il *leader* della logica esatta in Germania, segue Sigwart nell'opinione ricordata, andando se possibile ancor oltre... Si può osservare che nessuno si metterebbe a studiare logica se non avesse già stabilito che gli uomini sono tanto atti a errare nel loro senso di logicità quanto lo sono a ragionar male, se non sostenesse che la distinzione tra ragionar bene e ragionar male è che il primo conduce alla conoscenza della *verità*, mentre l'altro no, e che per verità si intende qualcosa che non dipende dal modo come noi pensiamo o sentiamo che essa sia». È chiara la netta opposizione a tutta l'impostazione psicologistica dell'algebra della logica europea – che troveremo appunto ribadita e codificata da Schröder – e nel contempo è evidente una consonanza, assai forte del resto nelle pagine di Peirce, con le tesi bolzaniana circa la natura «oggettiva» della logica secondo la quale essa si interessa alle «verità in sé», alle «proposizioni in sé», al pensabile piuttosto che al processo di pensare, tesi che troveremo decisamente riaffermata anche da Frege.

È tuttavia problematica e non univocamente determinata la risposta di Peirce circa il senso da attribuire a questa «oggettività» della logica. Ad essa concorrono tutta una serie di motivazioni che Peirce sviluppa in varie riprese e che sarebbe troppo lungo anche solo accennare in questa sede. Ci limitiamo a osservare come da un aspetto di queste motivazioni risulti inequivocabilmente la necessità, per Peirce, di un richiamo all'attività e all'espressione *simbolica* in generale: «L'ordito e la trama di ogni ricerca sono i simboli, e la vita del pensiero e della scienza è la vita inerente nei simboli; sicché è errato dire semplicemente che un buon linguaggio è *importante* per un buon pensiero; poiché esso ne è l'essenza.» Il che d'altronde è in armonia con la sua più generale concezione dell'uomo quale essenzialmente «elaboratore di segni» (*sign-maker*) e «lettore di segni» (*sign-reader*): «La parola o segno che l'uomo usa *è* l'uomo stesso», o, ancora: «Il mio linguaggio è la somma totale di me stesso.» Si comprende chiaramente come, in questo contesto generale, Peirce si opponga con decisione a una concezione secondo la quale la logica «... riguardi in primo luogo il pensiero inespresso e solo secondariamente il linguaggio», essendo per lui

la logica «la scienza delle leggi generali dei segni e specialmente dei simboli».

Francesco Barone si esprime come segue, a conclusione di un'analisi del pensiero logico di Peirce: «Come nel caso della polemica contro il logicismo anche nella questione della validità formale l'opera di Peirce offre "anticipazioni" sugli sviluppi futuri della ricerca. E di altre "anticipazioni"... essa è ricca anche sul piano più strettamente tecnico... Ma una prospettiva storica che guardasse solo alle "anticipazioni" rischierebbe di essere semplicistica e di idolatrare come definitive le concezioni cronologicamente ultime. L'attualità autentica del Peirce... è forse ancor più nella sua capacità di tener presenti tutti i temi proposti al pensiero moderno dal rinnovamento della logica formale. Il riconoscimento della natura scientifica di questa, la rinunzia al suo tradizionale inserimento nelle discipline filosofiche furono eventi decisivi per il pensiero contemporaneo: ma come per ogni svolta fondamentale della storia umana, le conseguenze non erano predeterminate; la rivoluzione poteva essere negata in nome della tradizione o spingersi nel vuoto negando alla tradizione ogni senso. Per l'equilibrio con cui seppe concordare il "vecchio" e il "nuovo" sul piano della ricerca "logica", Peirce è un "esempio", e non solo del passato».

Volgendoci ora all'analisi dell'opera di Schröder, avremo modo di sviluppare ulteriormente questa problematica. La dialettica e critica apertura di orizzonti e la vasta problematicità che abbiamo riscontrato in Peirce verrebbero però cercate invano in Schröder, e alcune delle osservazioni precedenti dovrebbero già aver fatto intuire l'impostazione più circoscritta e «limitata» di quest'ultimo, che pur seppe dare, per parte sua, concreti e importanti contributi allo sviluppo e alla sistemazione dell'algebra della logica ottocentesca.

Ernst Friedrich Wilhelm Karl Schröder (1841-1902) fu professore di matematica a Karlsruhe dal 1876 al 1902. A lui si debbono i maggiori risultati – dopo quelli del Peirce – nell'elaborazione dell'algebra della logica di questo periodo. Iniziò la sua attività in questo senso col volumetto del 1877 *Der Operationskreis des Logikkalkuls* (*L'ambito operativo del calcolo della logica*): in esso Schröder abbandonava il significato booleano (esclusivo) della somma in favore di quello di Jevons e poneva l'accento sulla «legge di dualità» che connette teoremi espressi in termini dell'operazione « + »

o di «+» e «1» con corrispondenti teoremi espressi in termini di prodotto logico «×» o di «×» e «0». Il sistema risultante è l'odierna algebra della logica, che venne perfezionato nelle già citate *Vorlesungen*, il cui primo volume è appunto dedicato all'algebra delle classi, il secondo all'algebra delle proposizioni e il terzo (che ha come sottotitolo *Algebra und Logik der Relative* [*Algebra e logica dei relativi*]) alla logica delle relazioni. Il secondo volume è in due tomi, il secondo dei quali venne pubblicato postumo da Eugen Müller nel 1905; sempre postumo e a cura di Müller venne infine pubblicato nel 1909-10 un *Abriss der Algebra der Logik* (*Compendio di algebra della logica*) in due parti, la prima dal titolo *Elementarlehre* (*Teoria elementare*), la seconda dal titolo *Aussagentheorie, Funktionen, Gleichungen und Ungleichungen* (*Teoria degli enunciati, funzioni, equazioni e disequazioni*).

Non ci impegneremo qui in una descrizione tecnica dell'opera fondamentale di Schröder perché, come è stato più volte detto, essa è sostanzialmente un'opera di sistemazione generale di una problematica che, oggi almeno, si ritiene ormai «pragmaticamente superata» e da questo punto di vista ha in effetti solo un valore «storico» per quanto riguarda la conduzione tecnica della sistemazione stessa pur se estremamente ricca di stimoli e di suggerimenti, come hanno mostrato i ripetuti tentativi di Tarski di sistemazione della teoria delle relazioni. Ci limiteremo ad alcune considerazioni generali, tentando soprattutto di mettere in luce quelle che sono le motivazioni e le concezioni di fondo che a parere di Schröder giustificavano appunto l'intrapresa di un'opera così impegnativa e «completa», che sotto questo rispetto è tutt'oggi un documento unico per la comprensione di questa fase di sviluppo della logica.

Per Schröder la logica è un'indagine sulle «regole seguendo le quali si ricerca la conoscenza della verità»; d'altra parte la comprensione della verità è un atto di pensiero, sicché oggetto della logica in generale è «il *pensiero* in quanto ha come suo fine ultimo il *conoscere*». Ritornano subito alla mente le parole di Peirce in proposito: ognuno vede che si tratta di due concezioni di fondo diametralmente opposte. Sappiamo già, dal citato passo di Peirce, che Schröder mutuava questa sua concezione da Sigwart e tratta appunto da Sigwart è la determinazione del «pensiero *consequenziale* o *logico*» come pensiero «... connesso per l'intelletto indagan-

te con la coscienza della immediata comprensione o evidenza, quando una "necessità mentale" ci costringe a compierlo con la convinzione dell'assoluta certezza». Schröder assume come ideale teoretico per la sua opera quello di «portare alla coscienza le leggi del pensiero consequenzialmente corretto, di dar loro un'espressione generale che sia la più semplice possibile, di ridurle al minor numero di principi o assiomi e, soprattutto di improntare tale pensiero ad una consapevole abilità».

Accanto a questo momento teoretico generale, è esplicitamente dichiarata da Schröder un'aspirazione verso la sistemazione migliore e definitiva della materia: «La teoria è a tal punto sviluppata e perfezionata che sembrano già raggiungibili un'esposizione e un ordinamento definitivi per una prima e fondamentale parte dell'intero edificio.» Non che con ciò egli pensi l'algebra della logica come disciplina esaurita e «chiusa in sé», al contrario, vede la possibilità di un «ulteriore illimitato sviluppo» della materia ma lo vede nell'ambito di questo modello perfetto, definitivo, in funzione del quale si giustifica la sua aspirazione alla «completezza» delle *Vorlesungen*, completezza che va intesa come esame e discussione di tutti i lavori dedicati all'argomento e comunque di ogni contributo che possa aggiungere elementi alla fissazione della base teorica definitiva sulla quale si innesteranno i nuovi, previsti e auspicati «ulteriori sviluppi». Questa aspirazione fa sì, da una parte, che le *Vorlesungen* risultino appunto un documento essenziale per la conoscenza esatta e comprensiva dello stato dell'algebra della logica alla fine dell'Ottocento; d'altra parte permette al loro autore di apportare anche notevoli contributi tecnici al calcolo, come pure di avanzare osservazioni acute e distinzioni di effettivo valore. Abbastanza paradossalmente, tuttavia, essa gli preclude la possibilità di sviluppare una visione critica *generale* della logica del suo periodo; è proprio questa incapacità che si manifesta, ad esempio, nell'atteggiamento già più volte ricordato nei riguardi del pensiero e dell'opera di Frege.

Sul piano strettamente più tecnico, il punto di partenza di Schröder è il calcolo delle classi, cui quello delle proposizioni viene ricondotto mediante l'aggiunta di un postulato di «bivalenza», così da ottenere ciò che noi oggi chiamiamo appunto un'algebra a due valori (l'algebra semplice di Boole **2**). Come si ricorderà il postulato in questione (non equazionale) è della forma «se $x \neq 1$ al-

lora $x = 0$ e se $x \neq 0$ allora $x = 1$» e vincola appunto l'universo a contenere due soli elementi: il vero (1) e il Falso (0). Non stupisce allora che sia stato proprio Schröder ad introdurre l'impiego pressoché sistematico di «tavole di verità» per la valutazione delle formule proposizionali di cui abbiamo avuto occasione di parlare a proposito di Peirce.

Il momento logico fondamentale quindi per Schröder, come per Boole, è il calcolo delle classi (*Gebietkalkül*); egli lo chiama «calcolo identico» (sicché si parlerà di somma identica e non di somma logica, ecc.) perché lo riguarda come una «disciplina ausiliaria che precede la logica... ed è di natura puramente matematica». Schröder, che accoglie anche la distinzione booleana fra proposizioni primarie e secondarie, coglie quelli che possiamo ormai chiamare i soliti difetti della sistemazione booleana: non interpretabilità di ogni passo della deduzione, impiego di operazioni inverse, uso della somma esclusiva; egli evita ovviamente questi difetti nella propria sistemazione, la quale presenta altri pregi specifici che essenzialmente consistono nell'introduzione esplicita della classe vuota, nell'assumere la sussunzione (inclusione) come relazione fondamentale, nell'impiego costante e sistematico della legge di dualità e infine, non ultimo, nell'esposizione dettagliata, minuziosa e precisa. In proposito tuttavia va detto che considerazioni formali e informali o intuitive sono spesso così intricabilmente legate da rendere la lettura non sempre agevole. Interessante è la discussione che Schröder conduce in questo contesto del concetto booleano di universo del discorso; egli rileva l'interpretazione onnicomprensiva da parte di Boole di questo concetto (si veda tuttavia la nota 15 a pagina 136) e sulla base di questa discussione, oltre a mostrare come una tale interpretazione conduca a contraddizioni, sviluppa tutta una serie di acute osservazioni nelle quali si può senz'altro cogliere una precisa anticipazione della teoria russelliana dei tipi di cui ci occuperemo a suo tempo, dal momento che – per evitare appunto tali contraddizioni – Schröder suggerisce una sorta di «gerarchizzazione» dell'universo del discorso. Altra nozione interessante introdotta da Schröder è la distinzione fra molteplicità coerenti e incoerenti, che vedremo verrà rilevata anche da Cantor. L'interesse nel caso di Schröder è dato dal fatto che l'introduzione di tale distinzione avviene indipendentemente dalla scoperta delle antinomie della teoria degli insiemi.

D'altra parte il momento dogmatico e unidirezionale della concezione schröderiana viene messo in chiara luce quando egli dà la sua versione della sillogistica classica. Dal momento che egli ha introdotto la classe vuota, è necessario fare una distinzione fra la portata esistenziale delle proposizioni particolari e di quelle universali (si ricordi quanto abbiamo detto a proposito di Venn); sappiamo già che questo comporta precise limitazioni alle inferenze immediate possibili ed è noto d'altra parte che con ciò viene a cadere la validità di alcuni modi «imperfetti» del sillogismo classico. Ora per Schröder ciò significa semplicemente che tali modi sono scorretti (*in assoluto*), dal momento che la logica «esatta» è assolutamente corretta e «delle conversioni della logica tradizionale solo la *conversio pura* è ammessa nella logica esatta». Ne viene in particolare che i sillogismi che ora risultano non validi in effetti non sono per lui neppure più dei sillogismi nel vero senso della parola, per il semplice motivo che essi hanno tre e non due premesse (ove la terza premessa è appunto costituita dall'affermazione esistenziale); con questa conclusione Schröder ricalca un noto argomento di Kant, cui d'altronde egli si richiama esplicitamente; ciò equivale, si badi bene, a riproporre una sorta di identificazione fra logica esatta e «logica naturale».

Per quanto riguarda il calcolo delle relazioni non c'è in effetti, nella prospettiva da noi qui assunta, molto da dire: la fondazione teoretica di questo calcolo è dovuta a Peirce – cosa che del resto lo stesso Schröder riconosce a più riprese e molto «serenamente» (si ricordi il passo di Lewis in proposito che abbiamo sopra riportato) –; ovviamente, come caratteristica generale, la materia viene presentata con sistematicità e «completezza» certamente non riscontrabili in Peirce e l'elaborazione del terzo volume delle *Vorlesungen*, ha fornito il contesto generale e la sistemazione simbolica per la ricerca logica almeno fino all'apparizione dei *Principia*. Né mancano d'altra parte numerosi apporti originali dello stesso Schröder, due dei quali vogliamo qui ricordare.

Il primo riguarda il tentativo di applicare la teoria delle relazioni in ambito specificamente matematico: Schröder dedica la nona lezione del terzo volume alla trattazione, da questo punto di vista, della teoria delle catene (e del processo di inferenza per induzione ad essa strettamente collegato) che Dedekind aveva introdotto nel citato *Essenza e significato dei numeri*. L'idea era ovviamente quella

di estendere il più possibile l'algoritmo relazionale alla matematica ma il risultato di questo esperimento non fu incoraggiante: ne risulta una trattazione estremamente complicata e poco perspicua che non offre alcun vantaggio rispetto alla presentazione «matematica» tradizionale di Dedekind. Peirce stesso espresse in proposito il dubbio che «il professor Schröder attribuisca un valore troppo alto a quest'algebra».

Il secondo è invece assai più interessante sia perché è inserito nella ricerca di condizioni di validità per una formula, vale a dire nel tentativo di costituzione di un metodo di decisione che permetta di stabilire appunto la validità «assoluta» (ossia indipendente dal dominio di interpretazione) di una formula del calcolo, sia perché – malgrado la proposta di Schröder in questo senso si sia rivelata insufficiente – essa ha costituito lo spunto per tutta una serie di ricerche. Nella sua sistemazione Schröder introduce alcune relazioni particolari, che egli chiama «moduli», per le quali stabilisce un «abaco» che permette appunto di decidere la validità di ogni espressione che sia espressa in termini di soli moduli. Definisce quindi il processo di «condensazione» di un'espressione, intendendo con ciò il trasformarla in un'altra espressione che non contenga i simboli Σ e Π (noi oggi diremmo, in forza della «traducibilità» della teoria delle relazioni nel calcolo dei predicati del primo ordine: priva di quantificatori). Un'espressione che ammetta questa trasformazione viene detta «condensabile». In forza dell'abaco dei moduli, risulta che un'espressione «condensata» è decidibile rispetto alla validità. Ora Schröder congettura (in effetti: afferma decisamente) che *ogni* espressione del calcolo delle relazioni sia condensabile sicché se ne otterrebbe automaticamente la decidibilità rispetto alla validità per ogni espressione di tale calcolo (per la ragione precedente, si avrebbe la decidibilità del calcolo dei predicati del primo ordine). Già Korselt nel 1907 dimostrò che la congettura di Schröder è insostenibile presentando un'intera classe di formule che non la soddisfano, ma la cosa per noi interessante è che proprio a partire da questa problematica e da questi spunti schröderiani, alcuni anni più tardi Leopold Löwenheim ottenne uno dei risultati fondamentali per l'odierna ricerca logica, sul quale ritorneremo nel capitolo V.

I risultati tecnici ottenuti da Schröder non si fermano qui e molto ci sarebbe da dire ad esempio sui suoi tentativi di indagare

la definibilità elementare della proprietà delle relazioni, tentativi che sarebbero stati poi proseguiti da Löwenheim e Korselt. Ma – ricordato questo – ci sembra opportuno concludere con alcune osservazioni di Barone, il quale, valutando complessivamente il lavoro di Schröder, afferma che fu «... grave per una valutazione generale dell'algebra della logica... l'impressione suscitata dallo scompenso fra l'elaborazione della sua parte tecnica e la povertà della sua interpretazione filosofica. O l'attenzione veniva in tal modo attratta dalla trattazione scientifica, abbandonando come inutile l'inquadramento filosofico, o si era spinti ad accomunare nel disinteresse la scricchiolante fondazione dell'edificio su essa sorgente. Ciò che la nuova trattazione simbolica della logica richiedeva era una riflessione filosofica sul simbolismo che non raccogliesse dall'esterno frammenti di filosofia precostituita, bensì affrontasse originalmente anche i problemi tradizionali. Fu il lavoro che Peirce compì mirabilmente: e se alla fine del secolo scorso la frammentarietà della sua opera poteva farla passare in secondo piano rispetto alla "sistematicità" di Schröder, oggi è possibile apprezzarla in tutta la sua suggestività». Aggiungeremo solo che come elemento di paragone «positivo» potremmo assumere Frege al posto di Peirce, perché il discorso si attaglia perfettamente all'originalità della sua analisi e delle sue soluzioni a classici problemi di filosofia della matematica e della logica.

7. DALL'ALGEBRA DELLA LOGICA ALLA LOGICA ALGEBRICA

Come abbiamo più volte avuto occasione di sottolineare, l'algebra della logica ebbe un ruolo centrale in quel processo di progressivo sganciamento dai programmi filosofici tradizionali che caratterizza la logica matematica quale noi oggi la concepiamo. Sarebbe però un grave errore vedere in essa la tappa iniziale di uno sviluppo continuo ed ininterrotto che ha portato alla logica matematica contemporanea mantenendo inalterati obiettivi e metodi. Anche se Russell non esitò ad affermare che «la matematica pura fu scoperta da G. Boole nella sua opera pubblicata nel 1854», fu proprio la pubblicazione dei *Principia Mathematica* e la ripresa di temi e metodi fregeani inaugurata da Russell stesso e

da Whitehead ad aprire nuovi orizzonti, alternativi a quelli della logica algebrica.

Non è azzardato infatti affermare che il punto di partenza dell'indagine logica come oggi è praticata va ricercato, più che nei lavori degli algebristi, negli interventi di Frege e poi di Russell e Whitehead il cui obiettivo era di portare la logica nel contesto della problematica relativa ai fondamenti della matematica. È in Frege che troviamo la prima presentazione di un compiuto *linguaggio formale* in grado di esprimere teorie matematiche non banali, la prima adeguata esplicitazione di regole d'inferenza riguardanti tanto l'aspetto proposizionale quanto quello predicativo ed è in Frege infine che troviamo formulato in modo netto il progetto di un'indagine logica volta allo studio della struttura della matematica e alla sua fondazione. Come avremo occasione di vedere più in dettaglio, fu proprio quest'ultima esigenza che portò Frege a ricercare un linguaggio formale in cui *tutte* le dimostrazioni matematiche riguardanti i numeri naturali fossero riproducibili e quindi a chiedersi quali fossero gli ingredienti fondamentali nella struttura logica dei giudizi matematici. In questo modo Frege giunse a mostrare *in concreto* come la logica delle proposizioni e quella dei termini fossero inadeguate e si rendesse necessario partire da un'analisi più ravvicinata del linguaggio matematico rispetto a quella che – fatte le opportune precisazioni – l'algebra della logica accettava ancora nella sua sostanza. L'enfasi si spostava così ancora una volta dalle forme astratte dell'algebra all'analisi del linguaggio ed è questo aspetto che segna la differenza più profonda tra l'approccio di Frege e quello degli algebristi.

Storicamente – almeno fino agli anni trenta del nostro secolo – fu la via di Frege, di Russell e di Whitehead che lo sviluppo della logica seguì, ma lungi dall'affermare che *questa* è la logica e che il lavoro degli algebristi è stato solo un punto d'inizio, un primo passo poi superato, riteniamo che abbia senso parlare piuttosto di *due* tradizioni nello sviluppo della logica sino a quel periodo: una prima tradizione che ha origine appunto con Boole e che concepisce l'indagine come studio di diversi universi di proposizioni né più né meno come una qualsiasi teoria matematica astratta e in cui l'enfasi è posta sui rapporti tra le diverse strutture, i diversi universi, e un'altra tradizione che ha origine in Frege, che vede essenzialmente la logica come strumento, come costruzione di calcoli e si-

stemi deduttivi mediante i quali analizzare la struttura delle teorie, siano poi queste viste come immerse in un'unica Grande Logica (come appunto in Frege e in Russell) o considerate isolatamente.

Se come abbiamo già ricordato è stata questa seconda tradizione a modellare l'*immagine* attuale della disciplina logica così che una storia della logica, quale quella che stiamo conducendo, non può che seguire essenzialmente il suo sviluppo, è altrettanto necessario sottolineare come l'interscambio tra le due prospettive sia stato essenziale e sia sfociato – con la creazione della *logica algebrica* negli anni trenta – in una visione più ricca e articolata della ricerca logica nel suo complesso. La logica algebrica è diventata in questo modo un ramo della logica in cui non solo la logica delle proposizioni e quella delle classi, ma l'intera logica del primo ordine, quella d'ordine superiore, ecc. vengono analizzate con i metodi dell'algebra astratta. Non si tratta di una variante algebrica dell'approccio usuale, ma di una prospettiva che ha come obiettivo quello di porre in luce l'*unità* delle strutture concettuali al fondo del pensiero astratto e quindi un orizzonte diverso ma altrettanto significativo della tradizione logicista. Non deve stupire così che anche prima degli anni trenta questo tipo di indagine abbia portato a isolare problemi e metodi di estremo interesse, che furono poi assorbiti dall'altra tradizione: pensiamo ad esempio ad alcuni dei primi risultati di teoria dei modelli quali i teoremi di Löwenheim e di Skolem, alle indagini sul metodo assiomatico della scuola statunitense ai primi anni del secolo, al sorgere della logica modale di Clarence Irving Lewis (1883-1964). Gli esempi si potrebbero moltiplicare, ma è soprattutto in questi ultimi anni – con lo svilupparsi della logica categoriale e della semantica funtoriale nel senso di F.W. Lawvere – che l'approccio algebrico ha mostrato tutta la sua potenza ponendo in luce come sulla base del concetto algebrico di *categoria* non solo sia possibile fornire analisi profonde delle logiche di ogni ordine, ma trovare nuovi nessi tra logica e geometria, allargando il concetto stesso di universo di interpretazione.

Come ricordato nell'introduzione non ci è possibile nell'ambito della nostra storia fornire una trattazione adeguata di queste ricerche di straordinario interesse, per cui rimandiamo ai testi citati in bibliografia. Ci limiteremo a concludere questa parte dedicata all'algebra della logica con un *excursus* delle ricerche che dai lavori di Boole, Peirce e Schröder hanno portato agli sviluppi della logi-

ca algebrica. Questo ci porterà a fare riferimento ad argomenti non ancora trattati, ma il lettore non avrà difficoltà a ritornare eventualmente su queste pagine dopo aver preso visione dei capitoli successivi. D'altra parte – e questo spiega la lunghezza di questo paragrafo – abbiamo cercato di rendere autonoma la trattazione almeno per quanto riguarda l'aspetto algebrico e matematico fornendo nei casi più importanti le definizioni ed illustrandole con esempi. Questo, tenuto conto del fatto che alcune almeno di queste nozioni saranno utili al lettore anche nel prosieguo dell'opera.

7.1 *Algebra della logica o calcoli logicisti?*

Parlando delle ricerche di Boole, Peirce e Schröder abbiamo più volte avuto occasione di sottolineare come questi studiosi ponessero l'attenzione, più che sul concetto di struttura astratta come oggi è d'uso comune, sulle nozioni di *ragionamento simbolico* e di *forma*. In questa prospettiva, le idee dominanti rimanevano da una parte quella di *sistema di regole*, di *calcolo*, dall'altra quella di *interpretazione* o di *applicazione*, spesso confuse, e solo in sordina si affacciava il sospetto che si dovesse distinguere tra aspetto *sintattico*, formale (il calcolo) e aspetto *semantico* (le interpretazioni) sottolineando che le interpretazioni devono avere uno specifico *dominio* e che differenze centrali tra interpretazioni diverse sono esplicitabili *solo* in riferimento a questi domini (come succede ad esempio quando confrontiamo interi e razionali e affermiamo che nel secondo caso ma non nel primo un elemento del dominio possiede un inverso moltiplicativo). Ancora nel 1897, nel suo *A Treatise on Universal Algebra* (*Trattato di algebra universale*), Alfred North Whitehead (1861-1947) poteva parlare di diversi *calcoli* che hanno *applicazioni* logiche, geometriche, fisiche, ecc., senza con questo individuare in un concetto generale di struttura l'elemento comune a tutte queste «algebre» e porre come oggetto di indagine lo studio delle operazioni e relazioni su di esse definite. L'elemento unificante rimaneva sempre il concetto di sistema di regole ed era sulla base di questa nozione che le varie algebre venivano introdotte e classificate, oscurando così la dualità tra aspetto semantico e sintattico.

Ciò non toglie che sia stato proprio Whitehead nel suo *Treatise* a formulare per la prima volta un sistema d'assiomi per il concetto

di algebra di Boole, cambiando le operazioni di base rispetto a quelle usate originariamente dallo stesso Boole ed in particolare sostituendo alla somma disgiunta la riunione. Sarà da questa formulazione assiomatica che prenderanno le mosse le indagini successive nel tentativo di analizzare più da vicino la struttura algebrica del calcolo introdotto da Boole. In termini più vicini ai nostri, possiamo dire che per Whitehead un'*algebra di Boole* è data da una struttura, una «interpretazione»

$$\mathbf{A} = < A, \wedge, \vee, -, 0, 1 >$$

in cui A è un insieme non vuoto, 0, 1 sono elementi di A, $\wedge$, $\vee$ e $-$ operazioni rispettivamente binarie (a due argomenti) e unaria (ad un argomento) che soddisfano le leggi seguenti, gli assiomi della teoria delle algebre di Boole:

A1a	$x \wedge x = x$	A1b	$x \vee x = x$
A2a	$x \wedge y = y \wedge x$	A2b	$x \vee y = y \vee x$
A3a	$x \wedge (y \wedge z) = $ $= (x \wedge y) \wedge z$	A3b	$x \wedge (y \vee z) = $ $= (x \vee y) \vee z$
A4a	$x \wedge (x \vee y) = x$	A4b	$x \vee (x \wedge y) = x$
A5a	$x \wedge (y \vee z) = $ $= (x \wedge y) \vee (x \wedge z)$	A5b	$x \vee (y \wedge z) = $ $= (x \vee y) \wedge (x \vee z)$
A6a	$x \wedge 1 = x$	A6b	$x \vee 0 = x$
A7a	$x \wedge 0 = 0$	A7b	$x \vee 1 = 1$
A8a	$x \wedge - x = 0$	A8b	$x \vee - x = 1.$

Se confrontiamo questi assiomi con quelli che abbiamo visto parlando di Boole notiamo subito alcune differenze che si possono riassumere in un'unica fondamentale: per Whitehead, il concetto di algebra di Boole si ottiene non già estendendo le regole di computo sui razionali e imponendo la legge di bivalenza, ma per astrazione dal modello concreto fornito dai *campi d'insiemi* in cui si ha una famiglia $\mathscr{F}$ di sottoinsiemi di un dato insieme I, contenente I stesso e l'insieme vuoto $\emptyset$, che risulta *chiusa* rispetto ad intersezione ($\cap$), riunione ($\cup$) e complemento ($\sim$). È questa scelta che permette a Whitehead di porre in luce la perfetta dualità tra $\cap$ e $\cup$ (che non esiste fra $+$ e $\times$).

Gli assiomi di algebra di Boole riflettono le proprietà algebriche di queste operazioni ove si vedano 0 e 1 rispettivamente come nomi di $\emptyset$ e I e $\wedge$, $\vee$, $-$ vengano interpretati come intersezione, riunione e complemento insiemistici. In questo modo scompare l'immediatezza dell'analogia con l'algebra dei razionali, ma emerge in modo netto la corrispondenza con le operazioni insiemistiche da una parte e i connettivi proposizionali dall'altra; basta osservare che se X e Y sono visti come le *estensioni delle proprietà* P e Q nel senso che se, indicando con $\in$ l'appartenenza (cfr. nota 6, pag. 295)

$$X = \{a \in I \mid P(a)\} \qquad\qquad Y = \{a \in I \mid Q(a)\}$$

allora avremo pure

$$\begin{aligned} X \cap Y &= \{a \in I \mid P(a) \text{ e } Q(a)\} \\ X \cup Y &= \{a \in I \mid P(a) \text{ o } Q(a)\} \\ \sim X &= \{a \in I \mid \text{non } P(a)\}. \end{aligned}$$

In altre parole il modo in cui le proprietà composte (P e Q), (P o Q), (non-P) si ottengono da P e Q corrisponde al modo in cui gli insiemi $X \cap Y$, $X \cup Y$ e $\sim X$ si ottengono da X e Y e le leggi riguardanti *entrambi* questi tipi di composizione si riflettono nelle leggi che seguono dagli assiomi per le algebre di Boole. Da questo punto di vista, il collegamento tra la logica delle classi e la logica delle proposizioni stabilito da Boole si ottiene come caso particolare una volta che si vedano le proposizioni P, Q ecc. come attributi, proprietà, di casi (i «tempi» di cui parlava Boole) e si proceda come sopra. Gli elementi 0 e 1 corrispondono allora rispettivamente alla *falsità* (l'essere *falso* in ogni caso) e alla *verità* (il verificarsi in ogni caso). Abbandonando il modello delle algebre di mucchi che per Hailperin stava al fondo delle indagini di Boole, gli assiomi dati da Whitehead contemperano tanto l'esigenza di completa algebrizzazione che quella di maggiore aderenza alla interpretazione logica sottolineata da Jevons.

Se in Whitehead non è sempre chiarissima l'idea che le algebre di Boole sono strutture (nel senso oggi comune) che pur potendo essere diverse hanno in comune la proprietà di soddisfare gli assiomi dati, ciò diviene evidente con Edward V. Huntington (1874-1952) che nel 1904 in un lavoro dal titolo *Sets of independent postulates for the algebra of logic* (*Insiemi di postulati indipendenti per l'algebra*

della logica), proprio partendo dagli assiomi di Whitehead definisce *algebra di Boole* ogni classe K su cui sono definite tre operazioni $\wedge$, $\vee$, e $-$ e che contiene due elementi privilegiati 0 e 1, così da soddisfare gli assiomi:

A1a $x \wedge y = y \wedge x$	A1b $x \vee y = y \vee x$
A2a $x \wedge (y \vee z) =$ $= (x \wedge z) \vee (x \wedge z)$	A2b $x \vee (y \wedge z) =$ $= (x \vee y) \wedge (x \vee z)$
A3a $x \wedge 1 = x$	A3b $x \vee 0 = x$
A4a $x \wedge - x = 0$	A4b $x \vee - x = 1.$

Huntington dimostra che i suoi assiomi sono equivalenti a quelli di Whitehead ed indaga in particolare tre diversi sistemi d'assiomi la cui coerenza, indipendenza e reciproca equivalenza dimostra in dettaglio utilizzando come modelli opportune algebre di Boole. Di particolare interesse – in vista degli sviluppi che la teoria avrà successivamente – è l'attenzione che Huntington pone sulla possibilità di vedere le algebre di Boole come particolari *strutture ordinate.* Dopo aver dimostrato che, data un'algebra di Boole **A**, se definiamo la relazione $\leq$ ponendo

$$a \leq b \quad \text{se e solo se} \quad a \wedge b = a$$

o equivalentemente

$$a \leq b \quad \text{se e solo se} \quad a \vee b = b$$

otteniamo un *ordinamento* del dominio A di **A** (cioè una relazione riflessiva, antisimmetrica e transitiva) rispetto al quale $\wedge$ e $\vee$ sono l'*infimo* ed il *supremo* nel senso che, dati a,b in **A**:

i) $a \leq a \vee b$; $b \leq a \vee b$ i') $a \wedge b \leq a$; $a \wedge b \leq b$

ii) Se $a \leq x$, $b \leq x$ allora $a \vee b \leq x$
ii') se $x \leq a$, $x \leq b$ allora $x \leq a \wedge b$,

Huntington dà un'assiomatizzazione delle algebre di Boole che assume come primitiva una relazione d'ordine $\leq$ sul dominio A e postula l'esistenza nel dominio di un *massimo* 1 e di un *minimo* 0, di

infimo $(a \wedge b)$ e *supremo* $(a \vee b)$ per ogni coppia di elementi $a, b \in A$. Il complemento $-a$ per ogni elemento $a \in A$ viene infine garantito da un assioma che impone che se $a \in A$, esiste (e, banalmente, è unico) l'elemento $-a$ che è simultaneamente il massimo degli x per cui $x \wedge a = 0$ ed il minimo degli y per cui $y \vee a = 1$. Come dichiarato da Huntington, questo sistema d'assiomi risale a Schröder (e da lui a Peirce) ed ha il grande merito di porre in primo piano, invece della relazione di identità, la relazione d'ordine $\leq$ di significato logico immediato. Scrivere infatti che $a \leq b$, se interpretiamo a e b come proposizioni ci permette di affermare che b è conseguenza di a mentre – se interpretiamo a e b come insiemi e leggiamo $\leq$ come inclusione – possiamo esprimere più in generale che la proprietà di appartenere a b è conseguenza di quella di appartenere ad a. La relazione $\leq$ di Huntington corrisponde alla $-\!\!<$ di Peirce e all'implicazione di McColl, così che questa assiomatizzazione rivela la sostanziale equivalenza tra l'approccio *equazionale* booleano e quello di tipo *inferenziale* che – come si ricorderà – McColl aveva contrapposto come più fondamentale a quello basato sull'identità.

Non ci è possibile soffermarci su altri aspetti di questi collegamenti, in particolare sui tentativi di fornire un punto di partenza unitario tanto alla geometria che alla logica effettuati prima da A.B. Kempe nel suo lavoro *On the relation between the logical theory of classes and the geometrical theory of points* (*Sulla relazione tra la teoria logica delle classi e la teoria geometrica dei punti*) del 1890 e poi da Josiah Royce (1855-1916) nel 1905 in alcune ricerche in cui confluivano tanto le indagini assiomatiche di Oscar Veblen (1880-1960) sulla geometria (in particolare l'assunzione della relazione «stare fra» come relazione fondamentale) quanto le ricerche di Peirce, Huntington e Schröder sull'algebra della logica, quanto infine i primi tentativi di prendere atto della portata della teoria generale degli insiemi di Cantor. Lo sfondo su cui si muovono tutte queste indagini – alcuni aspetti delle quali trovarono formulazioni lucide e fortunate nell'opera di Louis Couturat (1868-1914), in particolare nella sua *Algèbre de la logique* del 1905 (tradotta in inglese nel 1914 e in polacco nel 1918) – è costituito da quell'ampio e articolato complesso di ricerche sulla fondazione assiomatica di specifiche teorie matematiche (geometria, aritmetica, teoria dei numeri complessi e dei numeri reali, teoria dei campi e degli anelli, teoria

dei gruppi, ecc.) che a partire dai lavori sui fondamenti della geometria di Hilbert e dei suoi predecessori, come da quelli della scuola italiana (in particolare Peano e Pieri) portarono alla creazione di una vera e propria scuola assiomatica statunitense di cui oltre a Huntington fecero parte matematici come H.E. Moore, A. Bernstein, R. Moore, L. Dickson, J.W. Young tra gli altri.

Sino agli anni trenta i rappresentanti di questa scuola portarono avanti un progetto di ricerca relativamente autonomo – per quanto riguarda la logica – e in alcuni casi dichiaratamente antagonista a quello rappresentato dalla tradizione che si rifaceva a Frege, Russell e Whitehead. L'idea al centro dell'indagine era quella di *sistema deduttivo* astratto che si contrapponeva a quella di *calcolo* introdotta da Frege. Un sistema deduttivo è una famiglia K di oggetti astratti (le proposizioni) entro la quale viene isolata una sottofamiglia T (le proposizioni *vere* nel sistema). Mentre la famiglia K è chiusa rispetto a date operazioni che corrispondono ai connettivi che nell'approccio fregeano da proposizioni semplici ci danno proposizioni composte, la sottofamiglia T è definita in termini di *proprietà di chiusura* che corrispondono alla scelta di assiomi e di regole deduttive dell'approccio logicista.

La differenza essenziale rispetto alla nozione logicista di teoria è che i sistemi deduttivi sono sostanzialmente algebre in cui viene isolato un sottoinsieme in termini puramente algebrici. In questa prospettiva teorie e interpretazioni sono sullo stesso livello, in quanto in tutti i casi abbiamo a che fare con strutture astratte di cui si studiano matematicamente i rapporti e non con proposizioni e regole deduttive. La differenza tra le due impostazioni risulta in modo drammatico se consideriamo la logica pura. Dal punto di vista assiomatico un sistema logico è un particolare sistema deduttivo, oggetto d'analisi metateorica come i gruppi, il piano euclideo, ecc., è un *campo logico* come propone di chiamarlo Huntington in analogia con i campi numerici. La logica che si usa per studiare i campi logici è quella naturale che non ha bisogno di codificazioni in termini di assiomi e regole formali. In questa prospettiva ha senso – come di fronte a ogni teoria assiomatica – chiedersi se date proprietà dei campi logici sono fra loro indipendenti, se esistono campi logici che soddisfano date proprietà (formulate da un sistema d'assiomi) ecc. Tutto ciò è inconcepibile dal punto di vista logicista. Come Russell e Whitehead affermano esplicita-

mente: «i metodi riconosciuti per provare l'indipendenza non sono applicabili senza riserve alle nozioni fondamentali», e quelle logiche sono nozioni fondamentali per eccellenza. Ciò significa che dal punto di vista logicista, che mira ad una fondazione *assoluta* dei metodi astratti, i principi logici sono dati per via intuitiva o sperimentalmente isolati, ma non possono essere oggetto d'indagine. È così tagliata alla radice ogni possibilità di ricerca metamatematica e non è un caso che le prime indagini sull'indipendenza degli assiomi dei *Principia Mathematica* e sulla definibilità dei connettivi siano state condotte in tempi diversi da H.M. Sheffer (1913), P. Bernays (1926), E. Post (1921) e P. Henle (1931), piuttosto tardi e al di fuori della tradizione logicista.

Non deve stupire infatti che la polemica dei sostenitori dell'algebra della logica contro il metodo logicista abbia portato, ancora negli anni trenta, A. Bernstein, Huntington e altri a cercare di analizzare lo stesso sistema dei *Principia* come teoria deduttiva, in modo da porne in luce la struttura profonda oscurata dal metodo ivi adottato. Come ebbe a scrivere Bernstein nel 1931 (in un suo lavoro dal titolo significativo *Whitehead and Russell's theory of deduction as a mathematical science* [*La teoria della deduzione di Whitehead e Russell come scienza matematica*]) «le due forme di teoria, quella logicista e quella algebrica, stabiliscono gli stessi fatti; ma mentre nella forma booleana le idee che sono fuori della teoria sono separate da quelle che vi appartengono, nei *Principia* questa separazione non viene fatta. È l'impossibilità di distinguere tra questi due insiemi di idee che fa dire agli autori che "i metodi riconosciuti per provare l'indipendenza non sono applicabili senza riserve alle nozioni fondamentali". La teoria, una volta posta in forma booleana, non è più fondamentale di ogni altra teoria matematica ed è soggetta alle indagini postulazionali applicabili a tutte le altre scienze matematiche».

Va riconosciuto che i metodi che lo stesso Bernstein ed in generale tutta la scuola cui appartenne riuscirono ad elaborare per l'analisi metamatematica della logica non portarono a grandi innovazioni e si riducevano sostanzialmente da una parte ad un uso a volte raffinato delle *forme normali* che già Boole e Jevons avevano indagato, dall'altra ad un uso smaliziato delle *matrici logiche*, o *tavole di verità*. Il metodo – modellato sulle «tavole di moltiplicazione» per i gruppi date da A. Cayley (1821-1895) – consiste nel fissare

un dominio A di *valori di verità* (i possibili valori che le proposizioni possono assumere), isolarne un sottoinsieme, quello dei valori *designati* (nel caso classico in cui $A = \{ 0,1 \}$ il valore designato è solo 1, il vero) e nell'associare ad ogni connettivo n-ario f una tavola di k^n righe (se k sono i possibili valori di verità) che ad ogni possibile attribuzione di valori agli argomenti in $f(x_1,...,x_n)$ assegna il valore di $f(x_1,...,x_n)$ rispetto a quella attribuzione, dando così il *grafo* della funzione su cui viene interpretato il connettivo.

Così ad esempio ai connettivi $\wedge$, $\vee$, $\rightarrow$, e $\neg$ della logica classica sono associate le funzioni i cui grafi sono dati dalle tavole di verità:

x y	$x \wedge y$
1 1	1
1 0	0
0 1	0
0 0	0

x y	$x \vee y$
1 1	1
1 0	1
0 1	1
0 0	0

x y	$x \rightarrow y$
1 1	1
1 0	0
0 1	1
0 0	1

x	$\neg x$
1	0
0	1

Una volta fissato l'insieme dei valori di verità, quello dei valori designati e l'interpretazione dei connettivi, possiamo isolare la famiglia delle formule che assumono valori designati per *ogni* assegnazione di valore alle formule atomiche. Nel caso classico, sono queste le *tautologie* (il termine è di Wittgenstein) le leggi proposizionali classiche ma – come avremo occasione di vedere meglio parlando in capitoli successivi del sorgere delle logiche polivalenti – il metodo si applica a situazioni più generali e ci permette di analizzare problemi di indipendenza, definibilità e decidibilità per sistemi proposizionali in generale. Usato originariamente come procedura di verificazione da Peirce (in *On the Algebra of Logic* del 1885) e da Schröder, il metodo divenne lo strumento principale nelle prime indagini sull'indipendenza e sulla definibilità dei connettivi di Sheffer, Post e Bernays che abbiamo ricordato sopra. Vedremo più avanti, parlando della scuola polacca, come la tecnica delle matrici abbia avuto un ruolo fondamentale nello studio della logica proposizionale una volta che da metodo di verifica si trasformò in strumento algebrico.

Quello che ci preme sottolineare è che se – limitandosi alla logica classica ed alla logica proposizionale – la tecnica sembra adeguata a diverse questioni, i problemi iniziano quando dal piano

proposizionale si passa a quello predicativo e si ha a che fare con i quantificatori. Qui non bastano più le tavole di verità per definire e decidere quando una formula è legge logica o ha un modello, e si deve ricorrere a generalizzazioni infinitarie di supremi e infimi, congiunzioni e disgiunzioni infinite, che presentano nuove difficoltà sia che si vogliano trovare forme normali (e in questo caso si deve affrontare lo scoglio della distributività infinita su cui si era già arenato Schröder) sia che si vogliano elaborare metodi di decisione che – come le tavole – ci permettano di stabilire quando una formula è soddisfacibile o meno. Su questi temi la strada era stata aperta dalle indagini di Peirce e soprattutto dalla *summa* di Schröder che per lungo tempo costituì una miniera inesauribile. Fondamentali in questo senso furono nel primo decennio del nostro secolo i risultati ottenuti prima da Löwenheim e poi da Skolem su cui avremo occasione di ritornare più ampiamente in seguito; Löwenheim – utilizzando opportune forme normali – riuscì a ridurre il problema di decidere quando una formula è *soddisfacibile*, ha cioè almeno un modello che la rende valida, a quello di determinare quando ne ha uno numerabile. Questo, ricorrendo a sviluppi delle formule infinite su domini numerabili. Il risultato era profondamente significativo in quanto mostrava come – utilizzando quantificatori e congiunzioni o disgiunzioni infinitarie – il problema di decidere la validità logica o meno si riduceva al caso numerabile approssimando così quanto succede nel caso proposizionale in cui basta controllare un numero finito di assegnazioni per determinare se una formula è tautologia o meno. Come vedremo più avanti – e questo sarà uno dei grandi risultati degli anni trenta – il salto dall'infinito (anche se numerabile) al finito non è possibile e non esiste in generale procedura di decisione per formule che contengono quantificatori. Ciò non toglie che il risultato di Löwenheim provava la fecondità delle forme normali tanto nel caso proposizionale, finitario, che in quello infinitario in cui si analizzano i quantificatori come congiunzioni e disgiunzioni.

È in questa prospettiva che vanno viste diverse ricerche di Sheffer, Bernstein e soprattutto di Cooper Harold Langford (1895-1965) che sarà uno dei collaboratori di Lewis nella creazione della logica modale. Partendo dal concetto di forma normale (prenessa, disgiuntiva, congiuntiva, ecc.) questi studiosi cercarono sin dai primi anni venti di elaborare una vera e propria *metamatematica ge-*

nerale vedendo nei metodi di origine algebrica una via per affrontare questioni di coerenza, completezza, indipendenza, ecc. delle teorie. I risultati di Langford sugli ordini densi – anche se formulati nel linguaggio dei *Principia* – fanno riferimento essenziale a questo tipo di prospettiva ed assieme ai teoremi di Löwenheim e Skolem costituiscono il primo esempio di risultati significativi di quella che sarà poi la teoria dei modelli.

Utilizzando, come faceva Langford, le tecniche di eliminazione dei quantificatori introdotte da Skolem nel 1919 (non a caso all'interno del contesto algebrico) era possibile ricondurre il problema dell'esistenza di modelli all'analisi di esempi di opportune forme disgiuntive e quindi in ambito proposizionale booleano. L'algebra della logica permetteva di elaborare criteri per la soddisfacibilità di queste forme che consentivano di analizzare questioni di soddisfacibilità (e quindi di indipendenza, di coerenza, ecc.) *senza* costruire esplicitamente dei modelli e quindi senza doversi compromettere col problema di dove questi modelli andavano cercati, semplicemente sulla base della struttura di specifiche forme normali. In questo modo era possibile quella *system-form analysis* – per usare il termine introdotto da Sheffer – che costituiva un vero e proprio tentativo di metamatematica su base algebrica.

Il riferimento a Sheffer non è casuale perché è essenzialmente alla sua influenza che si devono molti degli sviluppi che la scuola assiomatica statunitense raggiunse in questo campo a partire dai primi anni venti e a cui accenneremo più avanti (cap. III, par. 5) parlando di Hilbert e dell'assiomatica dei primi anni del '900. È Sheffer che nel modo più lucido tematizzò la differenza di impostazione tra approccio algebrico-assiomatico e approccio logicista, sottolineando come mentre il primo parte da una molteplicità di teorie e ammette vari universi di interpretazione, il secondo mira alla costruzione di un'unica teoria (la Grande Logica) ed un unico universo di interpretazione. È ancora Sheffer che sottolinea l'importanza di una visione astratta dei sistemi formali, la necessità di tenere in conto la molteplicità delle diverse interpretazioni possibili e ne tematizza lo studio come oggetto centrale dell'indagine logica. In quest'ottica, ancora, non è strano che sia nel suo lavoro del 1915 (*A set of five independent postulates for boolean algebras, with application to logical constants* [*Un insieme di cinque postulati indipendenti per le algebre di Boole, con applicazioni alle costanti logiche*]) che per la pri-

ma volta compare l'espressione «boolean algebras» al *plurale*, con la esplicita osservazione che – diversamente da quanto lasciato intendere da Schröder, Whitehead, ecc. – esistono più *algebre* di Boole e non diversi usi dello stesso calcolo.

Sheffer tentò in diversi modi di colmare la divaricazione che si andava sempre più creando tra le tradizioni algebrica e logicista, sottolineando da una parte la maggior ricchezza espressiva che i *Principia* offrivano, dall'altra l'importanza di uno studio analitico delle teorie che ponesse in primo piano le loro proprietà metateoriche. Questa prospettiva si concretizzò in un vasto progetto di classificazione delle teorie sulla base di diverse nozioni di completezza e indipendenza (primalità, completezza esistenziale, completezza analitica, indipendenza completa, ridondanza, ecc.) studiate negli anni venti da E.H. Moore, A. Church, Langford e altri e nel tentativo di formulare a livello filosofico generale questa prospettiva astratta ed algebrica. Interessanti da questo punto di vista (anche se largamente dimenticati) sono i lavori che a questi temi dedicarono P. Weiss nel 1928 (in due articoli dal titolo *The nature of Systems*, pubblicati su «The Monist») e, ormai negli anni trenta, L. Kattsoff (la serie *Postulational methods*, I, II, III, su «Philosophy of Science» nel 1938).

Ma è nelle ricerche di C.I. Lewis dedicate al metodo assiomatico che l'eredità algebrica trovò un nuovo terreno di confronto con la logica dei *Principia*. Al fondo del progetto di elaborare un calcolo per l'implicazione stretta – su cui ci soffermeremo più a lungo nel capitolo V – sta infatti l'idea che, conforme alla tradizione algebrica, lo scopo della logica è quello non di elaborare un calcolo che si ponga come unica Grande Logica, bensì quello di studiare i rapporti tra proposizioni più significativi dal punto di vista astratto. Quindi, non solo l'implicazione e la deducibilità, ma anche l'indipendenza, la coerenza, ecc.: le distinzioni modali hanno per Lewis il ruolo di permettere una trattazione *unitaria* di tutte queste proprietà dei sistemi assiomatici.

7.2 *La teoria dei reticoli e l'algebra universale*

È ancora questo stretto collegamento tra prospettiva metateorica ed approccio algebrico che troviamo se consideriamo la ricerca

degli anni trenta quando – come ricordato – la trattazione algebrica della logica diverrà un elemento importante nel costituirsi della metamatematica come scienza in sé. Sullo sfondo di questa svolta estremamente significativa nella storia della logica e su cui avremo occasione di ritornare più ampiamente in seguito, stavano i nuovi sviluppi che la teoria astratta degli insiemi aveva portato nel campo dell'Analisi, della geometria e dell'algebra, in particolare con la creazione della teoria della misura e dell'Analisi funzionale ed il primo configurarsi della topologia generale e dell'algebra astratta.

Se la scuola polacca di Leopoli e Varsavia – almeno dagli anni venti – si specializzò soprattutto nell'approccio insiemistico all'Analisi e alla topologia, l'algebra astratta trovò la sua sistemazione classica in Germania per opera di Emmy Noether (1882-1935) e della sua scuola, primo fra tutti Bartel van der Waerden la cui *Moderne Algebra* (*Algebra moderna*) del 1930 costituì per molto tempo il testo di riferimento obbligato in questo campo. E. Noether raccoglieva l'eredità di quell'*analisi concettuale* delle nozioni matematiche fondamentali di cui Dedekind – il suo maestro – era stato uno dei principali propugnatori, ed è proprio a Dedekind che dobbiamo rivolgerci se vogliamo trovare l'origine di uno dei concetti chiave nel processo di algebrizzazione della logica, il concetto di *reticolo*, che appunto a partire dagli anni trenta andò sempre più rivelando la sua importanza.

Abbiamo ricordato come già Huntington – sulla scorta di Schröder e di Peirce – avesse posto in evidenza la struttura d'ordine delle algebre di Boole, in particolare il fatto che ogni algebra di Boole è un insieme ordinato *limitato* (vale a dire con massimo 1 e minimo 0) in cui per ogni coppia di elementi a,b esiste l'*infimo* ($a \wedge b$) e il *supremo* ($a \vee b$). In termini moderni, strutture ordinate di questo tipo (non necessariamente limitate) si chiamano *reticoli* («lattice» in inglese, «Verband» in tedesco, «treille» in francese) e se anche il nome e l'idea di porli al centro di una indagine sistematica delle strutture ordinate risale al 1933 ed è dovuta al matematico statunitense Garrett Birkhoff, essi erano stati studiati – oltre che da Peirce e da Schröder – da Dedekind, che fu il primo a mostrare come queste particolari strutture occorressero anche al di fuori del contesto dell'algebra della logica.

In particolare, esse comparivano all'interno di quella teoria dei

moduli (o degli ideali di numeri algebrici) mediante la quale Dedekind progettava di affrontare il problema della divisibilità in campi numerici arbitrari, non solo sul dominio degli interi. Per dare un'idea della cosa, consideriamo i naturali N e la relazione di divisibilità. È chiaro che essa costituisce un ordinamento, come è anche ovvio che dati $a,b \in$ N, il supremo $a \vee b$ sarà il minimo comune multiplo e l'infimo $a \wedge b$ il massimo comun divisore dei due numeri (come Schröder ampiamente sottolineava nelle *Vorlesungen*). Abbiamo così un reticolo e la teoria della divisibilità ammette una traduzione in termini reticolari, nel senso che la struttura indotta dalla divisibilità su naturali si riflette in questo reticolo. Il discorso si può estendere a domini numerici arbitrari considerando *ideali* e – al posto della divisibilità – l'inclusione: anche in questo caso otterremo reticoli i cui elementi sono ideali e la cui relazione d'ordine è l'inclusione. Se poi concentriamo la nostra attenzione sulla *struttura di gruppo* degli ideali (gli ideali sono sottogruppi del dominio in cui si trovano) otteniamo che anche i sottogruppi di un gruppo costituiscono un reticolo, addirittura un reticolo *completo* nel senso che *ogni* famiglia $\{a_i\}$ di elementi del reticolo ha supremo ed infimo, non solo quelle di due elementi (o, il che è lo stesso, di un numero *finito* di elementi) come i reticoli qualsiasi.

Quello che Dedekind osservò è che i sottogruppi additivi di un anello e i sottogruppi normali di un gruppo godono di una particolare proprietà, la cosiddetta *legge modulare*:

$$\text{Se } x \leq z \text{ allora } x \vee (y \wedge z) = (x \vee y) \wedge z.$$

Questa legge è una conseguenza della distributività che abbiamo incontrato nella assiomatizzazione delle algebre di Boole data da Whitehead, per cui

$$\begin{aligned} a \wedge (b \vee c) &= (a \wedge b) \vee (a \wedge c) \\ a \vee (b \wedge c) &= (a \vee b) \wedge (a \vee c) \end{aligned}$$

ma si può provare che esistono reticoli modulari (in cui cioè vale la legge di modularità) e che pure non sono distributivi, cosicché la modularità è strettamente più debole della distributività. L'interesse dei reticoli modulari non sfuggì a Dedekind che ottenne i primi risultati significativi al riguardo, risultati particolarmente

interessanti in quanto – come Birkhoff e von Neumann posero in luce nel 1936 – reticoli di questo genere si presentano in modo naturale nello studio degli aspetti logici della meccanica quantistica (parleremo nell'ultimo capitolo della nostra storia di alcuni degli sviluppi di queste ricerche). Quel che è ora importante sottolineare è che il rapporto tra modularità e distributività portava in primo piano il problema delle condizioni che un reticolo deve soddisfare per essere distributivo. Già Schröder e Korselt, prima di Dedekind, avevano fornito esempi di reticoli non distributivi e non modulari e la cosa risultava importante dal punto di vista logico in quanto da una parte la distributività è importante per ottenere risultati di *forma normale* (disgiuntiva o congiuntiva) e quindi per valutare la cardinalità di un dato reticolo in funzione della cardinalità dei suoi generatori, dall'altra in quanto ci sono particolari rapporti tra esistenza di complementi e distributività, e quindi tra il concetto di reticolo e quello più speciale di algebra di Boole.

Dal punto di vista della teoria dei reticoli, un'algebra di Boole non è altro infatti (se consideriamo gli assiomi di Whitehead) che un reticolo limitato distributivo, *complementato*, tale cioè che per ogni elemento a esiste un elemento $-a$ tale che:

$$a \wedge -a = 0 \qquad\qquad a \vee -a = 1$$

La distributività ci consente di affermare che $-a$ sarà *unico* e quindi che ogni algebra di Boole **B** è chiusa rispetto a una funzione $- : \mathbf{B} \rightarrow \mathbf{B}$ che ad ogni elemento a associa *il* suo complemento $-a$. Strettamente legato a questa unicità è la caratterizzazione che abbiamo già visto in Huntington di $-a$ come il *massimo* x tale che $a \wedge x = 0$ è il *minimo* y per cui $a \vee y = 1$. Fu Peirce a porre la domanda se era possibile *caratterizzare* le algebre di Boole solo come reticoli limitati dotati di un unico complemento e provare così che all'inverso l'unicità del complemento comportava la distributività. Il lavoro di Huntington che abbiamo più volte citato costituisce il primo di numerosi tentativi di rispondere all'interrogativo, tentativi che si scontrarono tutti con la difficoltà di costruire reticoli con complemento unico non distributivi o di provare la loro non esistenza. Una risposta venne solo nel 1940 per opera di R.P. Dilworth che provò il sorprendente risultato che *ogni* reticolo (anche non distributivo) può essere immerso in un reticolo con com-

plemento unico, cosicché la congettura di Peirce ha risposta negativa. I problemi riguardanti modularità e distributività risultavano in questo modo strettamente connessi con questioni riguardanti la chiusura rispetto ad operazioni diverse dal passaggio all'infimo e al supremo e l'importanza di tali questioni non tardò ad emergere mano a mano che si andavano scoprendo sempre nuovi contesti in cui la nozione di reticolo risultava centrale.

Come abbiamo detto, fu Garrett Birkhoff che in un articolo dal titolo *On the combination of subalgebras* (*Sulla combinazione di sottoalgebre*) del 1935 presentò lo studio dei reticoli come una vera e propria teoria dotata di metodi e problemi propri, oltre che di significative applicazioni, e non più come una «ricerca di tipo assiomatico» come ancora Emmy Noether aveva visto i risultati di Dedekind sulla modularità. Dei metodi della teoria (soprattutto di quelli che ebbero rilievo nella logica algebrica) parleremo più avanti; quello che è opportuno sottolineare qui – rimandando a testi più adeguati per una trattazione meno sommaria di questi temi – è l'*ampiezza* dei tipi di collegamento tra la teoria dei reticoli ed il resto della matematica che sono andati via via emergendo: tutti questi collegamenti non avrebbero tardato a mostrare il loro significato logico generale una volta che fosse emersa la sostanziale intertraducibilità tra reticoli e sistemi proposizionali.

L'obiettivo dichiarato di Birkhoff era di fornire con la teoria dei reticoli uno strumento «per affrontare i problemi combinatori nell'algebra astratta», in particolare alla base di questa speranza era il fatto che tanto le sottoalgebre di un'algebra data che le sue congruenze forniscono dei reticoli completi in termini dei quali tentare una classificazione delle strutture. Il contesto era quindi l'algebra universale che appunto in quegli anni Birkhoff, Oystein Ore e altri andavano edificando, dando finalmente corpo al progetto, che Whitehead aveva formulato nel suo *Treatise*, di uno studio coordinato degli aspetti comuni di tutte le specifiche strutture che andavano emergendo: gruppi, anelli, campi, ideali, moduli, ecc. Era appunto questa ricchezza di esempi concreti e di contesti specifici che permise a Birkhoff, in un fondamentale lavoro del 1935, *On the structure of abstract algebras* (*Sulla struttura delle algebre astratte*) di formulare il programma di quell'algebra universale che sarebbe poi divenuta negli anni cinquanta uno dei fattori centrali nello sviluppo della teoria logica dei modelli. Contemporanei agli interventi di Birkhoff

furono quelli di Ore che in una serie di lavori (a partire da *On the foundations of Abstract Algebra* [*Sui fondamenti dell'algebra astratta*], anch'esso del 1935) generalizzò i classici risultati di Dedekind sui sottogruppi e sui moduli, mostrando come l'intera teoria si potesse formulare in astratto in termini di reticoli modulari.

Era quindi aperta la strada ad una teoria della decomposizione e della dimensione per strutture arbitrarie che soddisfacessero date condizioni molto generali (in particolare la *condizione della catena discendente* introdotta da E. Noether nello studio degli ideali di polinomi). In questo processo di generalizzazione a situazioni più ampie di risultati e nozioni algebriche nati in contesti particolari, i reticoli venivano a svolgere un ruolo fondamentale e – come ebbe ad affermare lo stesso Birkhoff nel 1966 – «risultava evidente che l'algebra universale avrebbe avuto ben scarse possibilità di esistere come disciplina non banale senza la teoria dei reticoli».

Ma le applicazioni della teoria non si riducevano a questo. Sempre attorno al 1935, più o meno in modo simultaneo, Hassler Whitney, lo stesso Birkhoff e Saunders Mac Lane introducevano i concetti di geometria combinatoria e di reticolo geometrico, e mostravano come essi comprendessero anche lo studio delle geometrie proiettive. Era così possibile uno studio astratto del concetto di dipendenza che includeva tanto l'analisi degli spazi vettoriali (e quindi concetti come base, dimensione, ecc.) quanto la dipendenza algebrica dei campi. Ci si incontrava in questo modo con altre nozioni chiave che in particolare Ore era andato analizzando: quella di *sistema* e di *operatore di chiusura* e quella di *connessione di Galois*, nozioni tra loro legate e di portata estremamente ampia che non avrebbero tardato ad avere anche applicazioni logiche.

Sul piano strettamente geometrico, le ricerche di Whitney e Birkhoff si incontravano invece con quelle di Karl Menger che – partito dal tentativo di dare una formulazione invariante alla geometria differenziale, indipendente da riferimenti a specifiche coordinate – era successivamente giunto a formulazioni generali del concetto di sistema di copertura e quindi al tentativo di esprimere la geometria proiettiva *n*-dimensionale *senza* riferimento ai punti, puramente in termini di sottospazi. Era qui che avevano un ruolo essenziale i reticoli ed avveniva l'incontro con le ricerche di Birkhoff e Whitney, rivelando un aspetto del concetto di reticolo che si

sarebbe rivelato di estremo interesse generale: la possibilità cioè di algebrizzare nozioni originariamente definite facendo riferimento alla distinzione tra punti e zone in termini del solo concetto di zona e quindi utilizzando solo enti fra loro omogenei. Era la tendenza che la topologia algebrica aveva mostrato essere estremamente feconda, sostituendo il concetto di figura come insieme di punti con quello di configurazione, di struttura combinatoria di figure (celle, simplessi, ecc.) componenti.

La stessa situazione – secondo Ore – si verificava nell'algebra, dove «nella discussione dei domini algebrici non si è interessati in prima istanza agli *elementi* [dei domini] ma alle relazioni tra certi *sottodomini particolari*» ed una situazione analoga era andata mostrando nell'Analisi Constantin Caratheodory a partire dal 1938, sviluppando una teoria dell'integrazione non più fondata sulla nozione di insieme di punti, ma su quella di *soma* (corpo o figura). In tutti questi casi le zone, le strutture, i corpi risultavano organizzati in reticoli più o meno arricchiti di altre operazioni (complementi, residuazioni, ecc.) e l'obiettivo era di studiare quanto ordinariamente espresso facendo riferimento ai punti, agli elementi, traducendolo in termini di reticoli.

Ciò naturalmente portava in primo piano un interrogativo che si sarebbe rivelato di importanza capitale: entro che limiti ogni reticolo di un dato tipo (modulare, distributivo, un'algebra di Boole, ecc.) è *rappresentabile* in specifici termini concreti? Ad esempio, è possibile dimostrare che ogni algebra di Boole è *isomorfa* ad un *campo d'insiemi* (cioè ad un'algebra i cui elementi sono tutti sottoinsiemi di un dato insieme I ed in cui $\wedge$, $\vee$, $-$, sono intersezione, riunione e complemento insiemistico)? Che cosa si può dire per i reticoli modulari, per quelli distributivi, ecc.? Solo una volta che si sia provato un risultato del genere è lecito dire che *ogni* asserto che possiamo dimostrare su campi d'insiemi o altre strutture concrete ricorrendo agli elementi in termini delle sole operazioni algebriche è già dimostrabile *senza* ricorrere alla nozione di elemento, utilizzando i soli assiomi per le algebre astratte. Un prototipo di teoremi di questo tipo (noti oggi come teoremi di rappresentazione) era già noto e si tratta del classico risultato di Cayley secondo il quale ogni gruppo è isomorfo ad un gruppo di permutazioni (o sostituzioni nel caso finito) su un insieme. Nel caso dei gruppi la dimostrazione è immediata e l'insieme cercato non è altro che il domi-

nio del gruppo stesso su cui si può immaginare agire ogni elemento moltiplicando a destra. La situazione non è così immediata nel caso delle algebre di Boole e dei reticoli in generale ed il problema è di come costruire il campo d'insiemi isomorfo all'algebra data. È chiaro infatti che non possiamo aspettarci che ogni algebra di Boole sia isomorfa ad un'*algebra delle parti* del tipo $\mathscr{B}(I)$ i cui elementi sono *tutti* i sottoinsiemi dell'insieme I. Ragioni di cardinalità ci dicono che algebre di questo tipo hanno cardinalità 2^k dove k è la cardinalità dell'insieme I e non tutte le algebre di Boole hanno una tale cardinalità.

Sino al 1936 il problema aveva ricevuto una risposta per le algebre di Boole e solo nei casi estremi: quello delle algebre finite e quello delle algebre complete e atomiche. In entrambi i casi, il concetto chiave era quello di *atomo*, introdotto a suo tempo da Schröder, che costituisce l'analogo algebrico – nel nostro contesto – di quello di punto o di elemento. Dato un reticolo $\mathbf{A}$ ed un suo elemento $a \in \mathbf{A}$, diciamo che a è un *atomo* se non è l'elemento minimo (nel caso questo esista) e se non esiste nel reticolo alcun elemento b, diverso anch'esso dal minimo, che sia strettamente minore di a, tale cioè che $b \leq a$ e $a \neq b$. Un reticolo si dice *atomico* se per *ogni* elemento a diverso dal minimo esiste almeno un atomo b per cui $b \leq a$; si dice invece *privo di atomi* nel caso non esistano atomi (cosicché ogni catena discendente è infinita). Ebbene, già Boole aveva sostanzialmente mostrato (utilizzando le forme normali) che ogni algebra di Boole finita ha cardinalità 2^n per qualche n finito ed è chiaro che ogni algebra di Boole, se finita, è atomica; se infatti, dato $a \neq 0$, non esistesse atomo $b \leq a$, potremmo avere una catena discendente infinita, contro la finitezza dell'algebra. Per ottenere la rappresentazione cercata, come Bernstein aveva mostrato, basta ora considerare la famiglia non vuota A degli atomi di $\mathbf{B}$ e prendere l'algebra delle parti $\mathscr{B}(A)$, che è ovviamente un campo d'insiemi. È immediato adesso provare che esiste un omomorfismo tra $\mathbf{B}$ e quest'algebra delle parti: basterà considerare la funzione

$$\rho\colon \mathbf{B} \rightarrow \mathscr{B}(A)$$

che ad ogni a associa l'insieme:

$$\rho(a) = \{b \in A \mid b \leq a\}$$

È facile verificare che

$$\rho(a \wedge c) = \rho(a) \cap \rho(c) \qquad\qquad \rho(a \vee b) = \rho(a) \cup \rho(b)$$
$$\rho(-a) = \sim\rho(a) \qquad \rho(0) = \emptyset \qquad \rho(1) = A.$$

Allo stesso modo si può verificare facilmente che non solo ρ è iniettiva, ma che sarà anche suriettiva (in quanto se $\{a_1,...,a_n\} \in \mathscr{B}(A)$, avremo che $\{a_1,...,a_n\} = \rho(a_1 \vee ... \vee a_n)$, dove $a_1,...,a_n \in \mathbf{B}$) ottenendo un isomorfismo tra **B** e il campo d'insiemi $\mathscr{B}(A)$.

Una tecnica non dissimile si poteva applicare – come mostrato da Tarski nel 1935 – anche al caso di **B** atomica e completa. Qui è la completezza che interviene per garantirci la suriettività e permetterci di affermare che se $\{x_i\}_{i \in I}$ è un elemento di $\mathscr{B}(A)$, allora

$$\rho\left(\bigvee_{i \in I} x_i\right) = \{x_i\}_{i \in I}$$

Il problema che rimaneva era quello delle algebre di Boole non complete (e quindi neanche finite) e non atomiche. Come sostituire gli atomi e ottenere analoghi degli elementi nel caso di algebre di Boole arbitrarie? La risposta venne nel 1936 ad opera di Marshall Stone con un famoso articolo, *The Theory of representations for Boolean Algebras* (*La teoria delle rappresentazioni per le algebre di Boole*) di cui è difficile sopravvalutare l'importanza sia dal punto di vista logico che da quello matematico generale. Stone infatti non solo forniva la soluzione al problema in esame, ma apriva la strada all'uso di metodi completamente nuovi applicabili – come ben presto si sarebbe visto – a classi molto varie di reticoli, non solo alle algebre di Boole. Facendo questo mostrava anche nuovi contatti con altre discipline matematiche, in particolar modo con la topologia generale. Come avrebbe scritto nel 1938, questo non era un fatto casuale: «Un principio cardinale della ricerca matematica moderna – scriveva Stone – si può concentrare nella massima: topologizza sempre», ed infatti la teoria della rappresentazione di Stone portava alla luce una dualità generale tra algebre di Boole e particolari spazi (i cosiddetti *spazi di Stone*, compatti, di Hausdorff e totalmente sconnessi) che avrebbero non solo gettato nuova luce su specifici problemi mostrando come fatti topologici si riflettono nelle proprietà delle algebre associate e viceversa, ma aprendo la strada a risultati analoghi per tipi più generali di reticoli che forni-

vano una nuova visione dell'intera teoria della rappresentazione per algebre. Se questo dà un'idea del significato particolarissimo del risultato di Stone per la matematica in generale, per apprezzarne la portata dal punto di vista logico è necessario esaminare più da vicino i concetti che Stone introdusse per provare il suo teorema.

L'idea di partenza costituiva in un certo senso un ritorno alle origini, in quanto Stone mostrava il senso profondo del collegamento fra campi numerici ed algebra della logica posto in luce da Boole, introducendo la nozione di *anello booleano*. Un anello si dice booleano se soddisfa la condizione di idempotenza

$$aa = a.$$

Si tratta della legge caratteristica che Boole aveva introdotto e Stone poteva mostrare che ogni anello booleano, oltre ad essere di necessità commutativo, dà origine ad un'algebra di Boole.

Utilizzando un nuovo sistema d'assiomi per le algebre di Boole introdotto da Huntington nel 1933, Stone poteva infatti mostrare come, dato un anello booleano $\mathbf{A}$, si ottenesse un'algebra di Boole $\mathbf{B}_\mathbf{A}$ definendo

$$-a = a + 1; \quad a \wedge b = ab; \quad a \vee b = a + b + ab$$

e come, inversamente, data un'algebra di Boole $\mathbf{B}$, si ottenesse un anello booleano $\mathbf{A}_\mathbf{B}$ ponendo:

$$a + b = (a \wedge -b) \vee (-a \wedge b);$$
$$ab = a \wedge b; \quad 1 = a \vee -a; \quad 0 = -1.$$

La corrispondenza risultava *biunivoca*, nel senso che si aveva tanto che l'algebra associata all'anello $\mathbf{A}_\mathbf{B}$ coincideva con l'algebra di partenza $\mathbf{B}$, quanto, all'inverso, che l'anello associato all'algebra $\mathbf{B}_\mathbf{A}$ coincideva con l'anello $\mathbf{A}$. Stone non mancava di sottolineare come questa idea di considerare unioni e intersezioni modulo 2 fosse un processo familiare ai topologi algebrici, ma il punto di importanza capitale era un altro: «l'identificazione delle algebre astratte che sorgono dalla logica e dalla teoria delle classi con sistemi trattabili con i metodi dell'algebra moderna» permetteva – come affermava Stone – di trasferire *in toto* alle algebre di

Boole tutta la strumentazione concettuale elaborata per lo studio degli anelli e in particolare l'intera teoria degli ideali. L'importanza di questo passo stava nel fatto che la teoria degli ideali era stata introdotta a suo tempo da Dedekind per soddisfare esigenze strettamente analoghe a quelle che si presentavano nella teoria della rappresentazione per algebre di Boole: come nella teoria degli anelli si trattava di generalizzare la teoria della divisibilità ed in particolare la nozione di numero primo, qui si trattava di generalizzare quella di atomo in modo da poter rappresentare ogni elemento di un'algebra di Boole come insieme, anche nel caso di algebre non atomiche.

Per anelli commutativi qualsiasi, un ideale D è un sottoinsieme di un anello $\mathbf{A}$ che risulta *chiuso* rispetto all'addizione (nel senso che se $a,b \in D$ anche $(a + b) \in D$) e tale che se $a \in D$ e b è un elemento qualsiasi dell'anello, $ab \in D$. L'esempio classico si ha considerando, dato un numero qualsiasi $a \in Z$, l'anello degli interi, la famiglia $[a]$ di tutti i multipli di a. È chiaro che se b, c sono multipli di a (cioè se $c = an$ e $b = ak$) avremo pure che $(b + c)$ sarà multiplo di a in quanto $(b + c) = (an + ak) = a(n + k)$. È pure chiaro che se n è un intero qualsiasi e b un multiplo di a, anche nb sarà un multiplo di a cosicché $[a]$ costituisce un ideale. Fatto importante, si verifica a/b se e solo se $[b] \subseteq [a]$, cosicché la relazione di divisibilità fra numeri si riduce a quella di inclusione fra ideali.

In Z tutti gli ideali hanno la forma $[a]$ per un a opportuno e sono quindi *principali*, generati cioè da un solo elemento. L'interesse del concetto di ideale nasce dal fatto che in anelli commutativi arbitrari possono esistere ideali che *non* sono di questa forma e quindi non sono individuati da un elemento dell'anello. Ciò permette di sviluppare una teoria della divisibilità in cui è possibile avere fattorizzazioni che non esistono nel senso ordinario, come prodotti di elementi. Per parlare di fattorizzazione di ideali è necessario definire una operazione di *prodotto* di ideali tra di essi ponendo che:

$$D \circ D' = [\{ab \mid a \in D, b \in D'\}]$$

dove la chiusura tra due parentesi [ ] indica l'ideale *generato* dall'insieme, cioè l'intersezione di tutti gli ideali che contengono l'insieme dato (e che, come è facile vedere, è un ideale). L'opportuni-

tà di questa nozione più lata di divisibilità sta ancora nel collegamento sopra ricordato, per cui

$$a/b \text{ se e solo se } [b] \subseteq [a]$$

ma il fatto interessante – che collega ideali e reticoli – è che la famiglia degli ideali in **A** ordinata dall'inclusione è un reticolo completo – come già ricordato, dove $\wedge$ corrisponde al massimo comune divisore e $\vee$ al minimo comune multiplo.

È naturale chiedersi a questo punto quali sono gli analoghi dei numeri primi in termini di ideali. La risposta è immediata: *primi* saranno quegli ideali che non si lasciano fattorizzare se non banalmente, quindi gli ideali *propri* D diversi da [1] tali che se $ab \in D$, o $a \in D$ o $b \in D$. È appunto questa la nozione che Stone poneva alla base della sua teoria della rappresentazione per le algebre di Boole, ma per capirne meglio il ruolo è opportuno vedere gli ideali non tanto negli anelli booleani, quanto nelle algebre di Boole associate e tradurre quindi le condizioni definitorie. Avremo così che un ideale (a questo punto la definizione è data per algebre di Boole o più in generale per reticoli) è un sottoinsieme D tale che:

I_1	Se $a, b \in D$	allora $a \vee b \in D$
I_2	Se $a \in D$ e $b \leq a$	allora $b \in D$.

Questa definizione è interessante perché ci permette di individuare contesti insiemistici diversi in cui gli ideali intervengono naturalmente. Di questo tema fondamentale negli anni dopo il 1935 si occuparono soprattutto i matematici polacchi, in particolare Alfred Tarski (1901-1983), Adolf Lindenbaum (1904-1941), Andrzej Mostowski (1913-1975) ed altri, che ottennero risultati estremamente interessanti al riguardo. L'idea di fondo è che il concetto di ideale è legato a quello di insieme «piccolo» se si assume che riunioni finite di insiemi piccoli è piccola e che se un insieme è piccolo ogni suo sottoinsieme lo è. Nozioni di «piccolezza» che soddisfano queste condizioni si trovano al cuore di diversi rami della matematica che proprio in quel periodo si andava sviluppando: nella teoria della misura (dove gli insiemi di misura nulla costituiscono un ideale) nella topologia generale (dove questo succede per

gli insiemi ovunque non densi o per quelli di prima categoria), ecc.; la stessa nozione di finito ci permette di individuare un altro tipo di ideale.

Tutti questi ideali si ritrovano entro algebre di Boole che – in ciascuna delle discipline citate – risultavano di grande importanza: le algebre di Boole degli insiemi misurabili secondo Lebesgue, degli insiemi boreliani, degli insiemi di misura finita, nel contesto della teoria della misura e dell'integrazione; nella topologia generale le algebre di insiemi simultaneamente chiusi e aperti (i cosiddetti «clopen») di uno spazio topologico, quella degli aperti regolari, ecc.; gli esempi si potrebbero moltiplicare. Queste algebre – come più in generale le famiglie di sottoinsiemi di un insieme chiuse rispetto ad opportune operazioni – vennero sistematicamente studiate da topologi come Kasimierz Kuratowski, analisti come Waclaw Sierpinski, logici ed insiemisti come Tarski, Lindenbaum, Mostowski, ecc.: l'introduzione del concetto di ideale risultava estremamente importante in questi contesti in quanto offriva nuovi metodi per costruire algebre di Boole con proprietà richieste.

Duale del concetto di ideale (definito su algebre di Boole) è quello di *filtro*, che per molti rispetti non avrebbe tardato a mostrarsi più significativo dal punto di vista logico e che risulta più intuitivo nella formulazione dei teoremi di rappresentazione. Esso fu formulato esplicitamente da Henri Cartan come risposta al problema di fornire una definizione di convergenza più generale di quella in termini di successioni e che fosse adeguata ai nuovi contesti che si presentavano nell'Analisi funzionale e nella topologia generale.

Un *filtro* (in un reticolo, o più specificatamente in un'algebra di Boole **B**) è un sottoinsieme F di **B** tale che

F_1 Se $a, b \in F$ allora $(a \wedge b) \in F$

F_2 Se $a \in F$ e $a \leq b$ allora $b \in F$.

Se consideriamo l'algebra delle parti $\mathscr{B}(X)$ dove $<X, \mathscr{O}>$ è uno spazio topologico, la famiglia degli *intorni* di un punto $x \in X$ (quei sottoinsiemi di X che contengono un aperto in cui si trova x) costituisce un filtro così che – come mostrava Cartan – si può genera-

lizzare il concetto di convergenza parlando di filtri che convergono a un punto. Osservando che la nozione stessa di topologia si può ricostruire sulla base del concetto di famiglia d'intorni assunto come primitivo, si vede come in questo modo la nozione di filtro venga ad assumere un ruolo capitale nella topologia generale; ruolo che soprattutto i matematici francesi legati a Nicolas Bourbaki avrebbero ampiamente sviluppato sia nel contesto della teoria degli spazi uniformi (con i lavori di André Weil e Paul Samuel) che in ambiti più specifici, ad esempio nella dimostrazione del teorema di Tykhonov secondo il quale prodotti di spazi compatti sono compatti, risultato questo che ha fondamentali applicazioni in campo logico.

I filtri non hanno una controparte naturale nella teoria degli anelli commutativi – dal cui modello partiva Stone – ma sono duali degli ideali nel senso che coincidono con gli ideali nel reticolo *opposto*, in cui si assume come relazione d'ordine la *conversa* della relazione d'ordine data nel reticolo di partenza. Passando da un reticolo $\mathbf{A}$ al suo opposto $\mathbf{A}^{\mathrm{P}}$, $\wedge$ diventa $\vee$ e $\vee$ diventa $\wedge$, il minimo 0 diventa il massimo e 1 diventa il minimo. Avremo così che – poiché tutti gli ideali contengono 0 e sono propri solo se non contengono 1 – tutti i filtri contengono 1 e sono propri solo se non contengono 0. In questa dualità – muovendosi entro algebre di Boole – agli ideali primi corrispondono gli *ultrafiltri*, cioè quei filtri (propri) per cui:

$$a \in F \text{ oppure } -a \in F$$

per ogni a appartenente all'algebra.

La situazione, come vedremo più avanti, si presenta nel caso di algebre di Boole ma non per reticoli più generali in cui duali degli ideali primi sono i filtri *primi*, tali cioè che

$$\text{se } a \vee b \in F \text{ allora } a \in F \text{ oppure } b \in F.$$

Nel caso delle algebre di Boole i due concetti coincidono in quanto per ogni $a \in \mathbf{A}$ abbiamo che $a \vee -a = 1$ e quindi $a \vee -a \in F$ mentre per reticoli più generali è il concetto di ultrafiltro il più forte. Esempi di ultrafiltri si trovano in abbondanza una volta che, in vista della loro dualità con gli ideali, si osservi che esiste

uno stretto rapporto tra ultrafiltri F e ideali primi D in un'algebra di Boole per cui si abbia che

$$a \in F \text{ se e solo se } -a \in D.$$

Così, se un ideale D contiene gli insiemi «piccoli» di un campo d'insiemi, il filtro F collegato conterrà gli insiemi «grandi» e viceversa. In particolare, se X è un insieme infinito e D l'ideale degli insiemi finiti, il filtro associato è noto come *filtro di Fréchet*, in quanto sono questi i filtri che si ottengono quando X è l'insieme dei naturali e si formula in termini di filtri la convergenza di successioni studiata appunto in generale da Maxim Fréchet nel 1906. Questo filtro conterrà la famiglia degli insiemi *co-finiti* (i cui complementi sono finiti). In modo analogo, è chiaro che gli insiemi misurabili dotati di misura non nulla formano un filtro, mentre un ultrafiltro è la famiglia di questi insiemi nel caso la misura assuma valori *solo* sull'insieme $\{0,1\}$. Misure di questo tipo risultavano particolarmente interessanti come mostravano i lavori che, proprio verso la metà degli anni trenta, Tarski, Lindenbaum e Stanislaw Ulam avevano dedicato al *problema generale della misura*, lavori che avrebbero portato Ulam a introdurre il concetto di cardinale *misurabile* (sul quale ritorneremo nel capitolo VII parlando della problematica legata ai grandi cardinali).

Gli esempi del ruolo decisivo di ultrafiltri e ideali primi si potrebbero moltiplicare, ma la domanda centrale rimaneva sempre quella della loro esistenza. Gli unici ultrafiltri non banali che – per così dire – sono dati in natura, sono quelli determinati da atomi, nel senso che hanno la forma

$$F = \{x \in B \mid a \leq x\}$$

dove a è un atomo. Di fatto – come Tarski aveva mostrato nel 1935 – in un'algebra di Boole completa e completamente distributiva (in cui cioè valgono le leggi distributive per $\wedge$ e $\vee$ su famiglie arbitrarie) questi sono gli *unici* ultrafiltri esistenti: l'algebra è atomica ed è isomorfa ad un'algebra delle parti (l'algebra $\mathscr{B}(A)$, dove A è la famiglia degli atomi). Per utilizzare gli ultrafiltri come generalizzazioni degli atomi era necessario allora *provare* l'esistenza di sufficienti ultrafiltri, senza ipotesi di completezza o atomicità. In mancanza di espliciti metodi effettivi di costruzione, la via più di-

retta che rimaneva era il ricorso a quell'assioma di scelta che Zermelo aveva introdotto nel 1904 e già utilizzato per dimostrare il teorema del buon ordinamento.

Seguendo una linea dimostrativa che sarebbe poi divenuta classica, Stone considerava allora – dato un elemento a ed un altro elemento b per cui $a \not\leq b$ – la famiglia $\mathcal{K}$ dei filtri nell'algebra che contengono a ma non b. Essa è ordinata dalla relazione di inclusione e ricorrendo a un procedimento per induzione transfinita come Zermelo, e che si basava sull'ipotesi di disporre di un buon ordinamento dell'algebra, Stone poteva provare che la famiglia possedeva un elemento *massimale* F, un filtro cioè che non risultava strettamente incluso in nessun altro filtro della famiglia. Si trattava a questo punto di provare che ogni filtro massimale è un ultrafiltro col che si aveva il teorema (noto appunto come *teorema dell'ultrafiltro* di Stone) in base al quale in ogni algebra di Boole $\mathbf{B}$, se a e b sono elementi per cui $a \not\leq b$ esiste un ultrafiltro F che contiene a ma non b.

Il risultato viene in generale formulato in questa forma anche se – come ricordato – basandosi sull'analogia con gli anelli, Stone dimostrò propriamente il teorema duale per cui, nelle condizioni di sopra, esiste un ideale primo D che contiene b ma non a. Per anelli commutativi generali si disponeva di un risultato d'esistenza per ideali primi già dal 1929, dimostrato da Wolfgang Krull nel contesto della sua teoria assiomatica degli ideali. D'altra parte già E. Noether aveva utilizzato l'assioma di Zermelo nella teoria generale degli ideali e Krull aveva fatto uso dello stesso principio per dimostrare che ogni anello commutativo si estende ad un anello algebricamente chiuso. L'uso dell'assioma di scelta in contesto algebrico per provare l'esistenza di oggetti massimali (applicazioni, sottostrutture, basi, ordinamenti, ideali, ecc.) nella teoria dei campi e degli anelli era stato introdotto da E. Steinitz, E. Artin e O. Schreier nei primi anni venti e diverrà col tempo – come avremo occasione di vedere più avanti – una costante. In quest'ottica, il risultato di Stone si presentava come una forma particolare del *teorema fondamentale dell'aritmetica degli ideali*, secondo il quale in ogni anello booleano ogni ideale diverso dall'unità [1] è il prodotto di ideali primi. Lo stesso Stone sottolineò questo aspetto, mentre d'altra parte sul versante delle algebre di Boole una versione del teorema dell'ultrafiltro per campi d'insiemi motivata da consi-

derazioni di teoria della misura era già stata data da Tarski nel 1930.

L'aspetto rivoluzionario per la logica algebrica come per la matematica in generale stava in un'altra direzione, nell'uso cioè che Stone faceva di questo teorema di esistenza per rappresentare come campo d'insiemi le algebre di Boole e nella dualità che istituiva tra algebre e spazi. Provato infatti che, dati due elementi distinti a e b di un'algebra $\mathbf{B}$, esiste un filtro F che contiene a e non contiene b, Stone passava a considerare l'algebra delle parti $\mathscr{B}(X)$, dove X è la famiglia degli ultrafiltri nell'algebra di Boole $\mathbf{B}$ e definiva l'applicazione ρ: $\mathbf{B} \rightarrow \mathscr{B}(X)$ ponendo:

$$\rho(a) = \{F \in X \mid a \in F\}.$$

Per il teorema dell'ultrafiltro, è chiaro che elementi distinti avranno immagini distinte e che quindi l'applicazione sarà *iniettiva*. D'altra parte dalla definizione di ultrafiltro è chiaro che sarà un *omomorfismo* e che conserverà le operazioni $\wedge$, $\vee$, $-$, cosicché:

$$\rho(a \wedge b) = \rho(a) \cap \rho(b); \qquad \rho(a \vee b) = \rho(a) \cup \rho(b)$$

$$\rho(0) = \emptyset \qquad \rho(-a) = \sim\rho(a) \qquad \rho(1) = X.$$

La necessità di considerare ultrafiltri e non semplicemente filtri risulta chiara se consideriamo l'equazione $\rho(-a) = \sim\rho(a)$ che esprime il fatto che l'applicazione ρ conserva il complemento. Per provare che $\sim \rho(a) \leq \rho(-a)$ occorre mostrare che se $F \notin \rho(a)$, cioè $a \notin F$, allora $-a \in F$ e questa non è altro che la condizione definitoria di ultrafiltro, mentre l'inclusione $\rho(a \vee b) \leq \rho(a) \cup \rho(b)$ non è che la controparte del fatto che i filtri sono primi (vale a dire che se $(a \vee b) \in F$ allora $a \in F$ oppure $b \in F$). Per reticoli più generali delle algebre di Boole i due concetti sono distinti ed è importante osservare – come avrebbe fatto lo stesso Stone in un lavoro successivo sul quale torneremo più avanti – che per la conservazione di $\vee$ è sufficiente la primalità. Quello che conta è che a questo punto per rappresentare come campo di insiemi l'algebra $\mathbf{B}$ ci basta osservare che la ρ sopra definita è un monomorfismo da $\mathbf{B}$ a $\mathscr{B}(X)$ che diventa un isomorfismo una volta che al posto di $\mathscr{B}(X)$ si consideri l'*immagine* di $\mathbf{B}$ rispetto a ρ, $\rho(\mathbf{B})$, i cui elementi

sono *tutti e soli* gli insiemi di forma $\rho(a)$ per $a \in B$. In quanto sottoalgebra di un'algebra delle parti, $\rho(\mathbf{B})$ sarà un *campo d'insiemi* e l'isomorfismo ci dà così la rappresentazione richiesta.

Come mostrava Stone però, si poteva dire di più e *caratterizzare* in astratto i campi d'insiemi che sono sufficienti per rappresentare *ogni* algebra di Boole. È a questo punto che interveniva l'elemento più nuovo delle indagini di Stone, la dualità tra algebre di Boole e spazi topologici. Sull'insieme X degli ultrafiltri di $\mathbf{B}$ possiamo infatti definire una *topologia di chiusi*, cioè una famiglia $\mathscr{C}$ di sottoinsiemi di X che contiene X e l'insieme vuoto $\emptyset$ e che è chiusa rispetto a intersezioni arbitrarie e riunioni finite. Famiglie di questo tipo – secondo la definizione fondamentale data da Hausdorff – costituiscono una topologia di chiusi per lo spazio $<X, \mathscr{C}>$, mentre dualmente topologia di aperti è ogni famiglia $\mathscr{O}$ che contiene sempre X e $\emptyset$ ma è chiusa rispetto a intersezioni finite e riunioni arbitrarie.

Per trasformare X in uno spazio topologico basterà dare una topologia di chiusi e prendere come elementi di $\mathscr{C}$ tutte le intersezioni (finite o infinite) di elementi del tipo $\rho(a)$ dove $a \in \mathbf{B}$. In questo modo la famiglia

$$\mathscr{C}_0 = \{\rho(a) \mid a \in \mathbf{B}\}$$

diviene una *base* della topologia così definita e gli elementi di $\mathscr{C}_0$ sono appunto i *chiusi della base* dello spazio. Che quella così definita sia effettivamente una topologia è di verifica immediata, in quanto $\mathscr{C}_0$ è chiuso rispetto alle riunioni finite e quindi lo stesso sarà per $\mathscr{C}$; la chiusura rispetto a intersezioni arbitrarie è ovvia per definizione. $\mathscr{C}_0$ è chiuso rispetto a riunioni finite in quanto $\mathbf{B}$ lo è rispetto a $\vee$; $\mathbf{B}$ è chiusa però anche rispetto a $\sim$ cosicché a sua volta $\mathscr{C}_0$ risulterà chiusa rispetto al complemento. Ciò significa – ricordando che gli aperti di uno spazio topologico coincidono con i complementi dei chiusi – che gli elementi di $\mathscr{C}_0$ sono anche aperti e che $\mathscr{C}_0$ è anche la base di una topologia di aperti (precisamente, la topologia i cui aperti coincidono con le riunioni arbitrarie di elementi di $\mathscr{C}_0$) topologia che dà lo stesso spazio che abbiamo definito sopra. Gli elementi di $\mathscr{C}$ sono quindi clopen.

L'algebra $\mathbf{B}$ è allora isomorfa ad un'algebra di clopen, ma di che tipo di spazio? Stone poteva mostrare due proprietà fondamentali degli spazi così costruiti a partire da algebre di Boole. La prima è

che si tratta di *spazi di Hausdorff* nel senso che, dati due punti (cioè due ultrafiltri) F e F' dello spazio esistono due aperti $\rho(a)$ e $\rho(b)$ *disgiunti* tali che $F \in \rho(a)$ e $F' \in \rho(b)$. Basta considerare qualsiasi elemento a tale che $a \in F$ e $a \notin F'$ (che dovrà per forza esistere essendo F e F' distinti) e prendere come b il complemento $-a$ per verificare il fatto. L'altra caratteristica è più profonda ed è la *compattezza*: lo spazio ottenuto, che possiamo indicare con $\mathbf{T_B}$, è *compatto* nel senso che ogni *copertura* di aperti dell'intero spazio X (cioè ogni famiglia $\{X_i\}_{i \in I}$ di aperti la cui riunione sia l'intero X) ammette un sottoricoprimento finito (nel senso che esiste un $J \subseteq I$, *finito*, per cui $\bigcup_{i \in J} X_i = X$). Dualmente, data una qualsiasi famiglia $\{Y_i\}_{i \in I}$ di chiusi essa ha intersezione vuota se ciò avviene per ogni sottofamiglia $\{Y\}_{i \in J}$ con J finito. Dalla compattezza – ricorrendo a un teorema dimostrato da J. Alexander – scende che ogni clopen dello spazio $\mathbf{T_B}$ appartiene a $\mathscr{C}_0$ cosicché, in conclusione, si può dire che ogni algebra di Boole $\mathbf{B}$ risulta isomorfa all'algebra dei clopen di uno spazio $\mathbf{T_B}$ compatto, di Hausdorff e totalmente sconnesso (vale a dire con una base di clopen).

Poiché è ovvio che ogni algebra di clopen è un'algebra di Boole (in quanto è chiusa rispetto a $\cap$, $\cup$ e $\sim$) si può provare che vale anche il viceversa e possiamo affermare che dato un qualsiasi spazio topologico $\mathbf{T}$ compatto, di Hausdorff, totalmente sconnesso (in breve, uno *spazio di Stone*) esisterà un'algebra di Boole $\mathbf{B}$ tale che

$$\mathbf{T} = \mathbf{T_B}.$$

In questo modo abbiamo una corrispondenza biunivoca fra algebre di Boole e spazi di Stone, corrispondenza che si sarebbe rivelata più tardi una vera e propria *dualità* tra categorie in quanto è facile provare che ad omomorfismi $f\colon \mathbf{B} \to \mathbf{B}'$ fra algebre di Boole è possibile associare applicazioni continue $t_f\colon \mathbf{T_{B'}} \to \mathbf{T_B}$ di direzione opposta tra gli spazi e viceversa. Dal punto di vista metodologico questo significava che da una parte è possibile studiare fatti topologici in termini di algebre di Boole, come sarebbe stato fatto ben presto dallo stesso Stone e da altri; dall'altra ciò comportava l'ingresso massiccio di metodi topologici nello studio dei reticoli e delle algebre di Boole.

Non è questa la sede per valutare sul piano matematico generale la portata di questi risultati; ci siamo soffermati così a lungo su di essi perché dal punto di vista dell'algebrizzazione della logica

hanno costituito una tappa decisiva e perché la quasi totalità dei metodi e dei problemi che successivamente sarebbero stati affrontati, in un modo o nell'altro si collegano ad essi. Questo è vero anche per le ricerche che parallelamente a Stone erano state condotte da Birkhoff, il quale già nel 1933 aveva ottenuto una rappresentazione essenzialmente equivalente a quella di Stone muovendosi nel contesto di quell'algebra universale che proprio in quegli anni si andava sviluppando e di cui egli era stato uno degli artefici principali. In questo ambito era stato possibile generalizzare concetti fondamentali come quelli di omomorfismo, isomorfismo, prodotto, quoziente, ecc. in modo da ottenere una teoria unificata per strutture algebriche (o algebre) qualunque. Astraendo dagli esempi dei gruppi, degli anelli, ecc., veniva definita «algebra» ogni struttura del tipo

$$\mathbf{A} = \langle A, \{f_i^{\mathbf{A}}\}_{i \in I} \{c_i^{\mathbf{A}}\}_{i \in J} \rangle$$

dove, per $i \in I$, $f_i^{\mathbf{A}}$ è una funzione, $f_i^{\mathbf{A}} : A^{n(i)} \rightarrow A$ a $n(i)$ argomenti e le $c_i^{\mathbf{A}}$, $i \in J$, sono elementi privilegiati di A, le cosiddette *costanti algebriche*. *Omomorfismo* da $\mathbf{A}$ a $\mathbf{B}$ diviene ogni applicazione $\mu : A \rightarrow B$ che conserva la struttura, nel senso che $\forall i \in I \ \rho(f_i^{\mathbf{A}} (a_1, \ldots, a_n) = f_i^{\mathbf{B}} (\rho a_1, \ldots, \rho a_n)$ e $\forall i \in I \ \rho(c_i^{\mathbf{A}}) = c_i^{\mathbf{B}}$; monomorfismi sono così gli omomorfismi iniettivi, isomorfismi quelli biiettivi. A questo livello di generalità era possibile unificare risultati e nozioni sorti originariamente nella teoria dei campi, degli anelli, dei gruppi, ecc., che furono ben presto applicati ai reticoli e alle algebre di Boole così come Stone aveva fatto per la teoria degli ideali.

Uno dei risultati centrali ottenuti da Birkhoff riguardava i prodotti sottodiretti, quelle sottostrutture dei prodotti $\Pi_{i \in I} \mathbf{A}_i$, ristrette alle quali tutte le proiezioni $\pi_i : \Pi_{i \in I} \mathbf{A}_i \rightarrow \mathbf{A}_i$ risultano suriettive.[34] Il problema generale che Birkhoff affrontava era da una parte quello di determinare le strutture sottodirettamente *irriducibili*

[34] Data una famiglia di strutture $\{\mathbf{A}_i\}_{i \in I}$, il prodotto $\Pi_{i \in I} \mathbf{A}_i$ è definito generalizzando quanto avviene nel caso finito. Dominio del prodotto sarà il prodotto cartesiano $\Pi_{i \in I} A_i$ costituito da tutte le funzioni $g : I \rightarrow \cup_{i \in I} A_i$ tali che ogni $i \in I$, $g(i) \in A_i$. Le operazioni vengono definite per componenti nel modo consueto. Si noti che nel caso I sia infinito, l'esistenza di almeno un $g \in \Pi_{i \in I} A_i$ richiede l'assioma di scelta, anzi è questa la formulazione «moltiplicativa» di tale assioma. Fissato $i \in I$, la funzione $\pi_i : \Pi_{i \in I} \mathbf{A}_i \rightarrow A_i$, definita ponendo $\pi_i(g) = g(i)$, è suriettiva e definisce un omomorfismo dal prodotto all'i-esimo fattore.

(tali cioè da non poter essere rappresentate come prodotti sottodiretti non banali) dall'altra quello di vedere entro che limiti era possibile rappresentare tutte le strutture di una data classe in termini di questi prodotti. Il risultato estremamente potente che Birkhoff aveva ottenuto è che *ogni struttura* algebrica è isomorfa ad un prodotto sottodiretto di strutture irriducibili e che – nel caso le algebre in esame costituiscano una classe *equazionale* (vale a dire siano tutte e sole le algebre che soddisfano un insieme di *equazioni*) – i fattori irriducibili della rappresentazione si possono trovare *entro* la classe stessa.

Utilizzando la sua caratterizzazione delle classi equazionali d'algebre come *varietà* (cioè come classi chiuse rispetto ai prodotti, al passaggio alle sottostrutture e ai quozienti) Birkhoff poteva concludere che una varietà è *determinata* dalla famiglia dei suoi elementi irriducibili. Tutti questi risultati – come la dimostrazione dell'esistenza di algebre libere per varietà – non facevano che porre in primo piano il ruolo della definibilità equazionale una volta che si generalizzasse da gruppi, anelli, ecc. ad algebre qualsiasi.

Come era noto almeno dalle prime assiomatizzazioni di Whitehead, le algebre di Boole *sono* una classe equazionale e lo stesso vale per i reticoli, i reticoli distributivi, ecc. Per vederlo basta isolare gli assiomi all'interno di quelli dati da Whitehead e lo stesso vale per tante altre strutture di questo tipo che mano a mano sarebbero emerse nell'analisi algebrica della logica e più in generale nella teoria degli insiemi ordinati. Se l'indagine della definibilità equazionale di classi di algebre è all'origine delle ricerche che a partire dal 1935 Tarski condusse sul calcolo dei sistemi e costituisce il paradigma di quella che poi sarebbe stata la teoria dei modelli, sul terreno più ristretto della teoria dei reticoli la verifica della definibilità equazionale comportava automaticamente la chiusura rispetto a prodotti, quozienti, sottostrutture per reticoli, algebre di Boole, ecc. con la possibilità quindi di applicare metodi generali dell'algebra universale.

È in questo contesto che il risultato di Birkhoff sui prodotti sottodiretti forniva una versione del teorema di rappresentazione per algebre di Boole. È facile vedere infatti che le algebre di Boole sottodirettamente irriducibili sono tutte isomorfe all'algebra *semplice* $\mathbf{2} = \langle\{0,1\}, \wedge, \vee, -, 0, 1\rangle$ dove $\wedge, \vee, -$ sono le funzioni definite dalle ordinarie tavole di verità a due valori, cosicché il teorema

viene a dire che per ogni algebra di Boole **B** esiste un monomorfismo $\rho : \mathbf{B} \rightarrow \mathbf{2}^I$ dove $\mathbf{2}^I$ è il prodotto di I copie dell'algebra **2**. La rappresentazione insiemistica di **B** si ottiene una volta che si vedano gli elementi del prodotto $\mathbf{2}^I$, che sono funzioni $f : I \rightarrow \{0,1\}$, come *funzioni caratteristiche* di insiemi X_f appartenenti all'algebra delle parti $\mathscr{B}(I)$. In questo modo il monomorfismo ρ ci permette di definire un monomorfismo $\overline{\rho} : \mathbf{B} \rightarrow \mathscr{B}(I)$ ponendo:

$$\overline{\rho}(a) = \{i \in I \mid \rho(a)(i) = 1\}.$$

Come le tecniche utilizzate da Stone che si basano sul concetto di filtro, anche quelle di Birkhoff avrebbero avuto notevoli applicazioni nella teoria dei reticoli e nell'algebrizzazione della logica. Quello che mancava loro era il collegamento con la topologia che ebbe storicamente un'importanza decisiva sul piano metodologico e che non ebbero le tecniche puramente algebriche di Birkhoff. L'elemento che nella rappresentazione di Birkhoff prendeva il posto dei filtri (e quindi del ricorso alla topologia spettrale) era dato dagli omomorfismi che collegano prodotti diretti e fattori e la sostanziale equivalenza delle due rappresentazioni nasceva dal fatto che – come Stone aveva notato nel suo lavoro – esiste tra omomorfismi, filtri e quozienti di reticoli distributivi un rapporto molto stretto che li rende sostanzialmente interscambiabili. La situazione generalizza quella che si verifica nella teoria degli anelli ed avrebbe avuto una grande importanza nel processo di algebrizzazione della logica dove gli omomorfismi costituiscono la controparte delle interpretazioni e vengono a loro volta individuati attraverso filtri opportuni.

Per meglio vedere la natura di questo rapporto, consideriamo due algebre di Boole **B** e **B**$'$ e supponiamo di avere un omomorfismo $h : \mathbf{B} \rightarrow \mathbf{B}'$. Associati ad h sono due sottoinsiemi di **B**, il *nucleo* Ker(h) e l'*antinucleo* Aker(h), definiti rispettivamente come le immagini inverse di 0 e di 1:

$$\mathrm{Ker}(h) = \{a \in \mathbf{B} \mid h(a) = 0\}\ ;\ \mathrm{Aker}(h) = \{a \in \mathbf{B} \mid h(a) = 1\}.$$

È immediato verificare che, quale che sia h, Ker(h) sarà un ideale e Aker(h) un filtro e la cosa vale in generale per reticoli distributivi con massimo e minimo; il fatto interessante è che si veri-

fica anche il converso, così che è possibile *caratterizzare* ideali e filtri come nuclei e antinuclei di omomorfismi.

Il punto di partenza è l'osservazione che – fissata l'algebra di Boole **B** – tutti gli ideali I ed i filtri F in **B** danno origine a relazioni di congruenza su **B** che indicheremo con $\equiv_I$ nel caso degli ideali e con $\equiv_F$ nel caso dei filtri. Basterà porre:

$$a \equiv_I b \qquad \text{se e solo se } (a \wedge -b) \vee (-a \wedge b) \in I$$

$$a \equiv_F b \qquad \text{se e solo se } (-a \vee b) \wedge (-b \wedge a) \in F.$$

Fatto questo, possiamo applicare la teoria generale delle congruenze su algebre sviluppata da Birkhoff e quozientare **B** in modo da ottenere $\mathbf{B}/I$ (nel caso si utilizzi la congruenza $\equiv_I$) o $\mathbf{B}/F$ (nel caso la congruenza sia $\equiv_F$): la procedura è definita in generale per congruenze qualsiasi. Data la congruenza $\equiv$ e l'elemento $a \in B$, indichiamo infatti con $[a] = \{b \in B \mid a \equiv b\}$ la classe di equivalenza di a rispetto a $\equiv$ e consideriamo l'*insieme quoziente*

$$B/\!\equiv\; = \{[a] \mid a \in B\}.$$

Quozientato il dominio B di **B**, si tratta ora di definire (secondo tecniche classiche dell'algebra universale) una struttura su $B/\!\equiv$ in modo che la *proiezione canonica* $\pi : B \longrightarrow B/\!\equiv$, che ad ogni a associa la sua classe di equivalenza $[a]$, sia un omomorfismo. Nel caso di algebre di Boole non è difficile verificare che basterà porre:

$$[a] \wedge [b] = [a \wedge b]; \quad [a] \vee [b] = [a \vee b]; \quad -[a] = [-a]$$
$$1 = [1]; \quad 0 = [0]$$

per avere che la proiezione canonica è un omomorfismo suriettivo il cui nucleo sarà $[0]$, mentre l'antinucleo coinciderà con $[1]$.

Nel caso la congruenza $\equiv$ sia determinata dall'ideale I, $\mathrm{Ker}(\pi)$ non sarà altro che l'ideale stesso e la medesima cosa si verificherà per F considerando invece l'antinucleo. Di più, come dimostrato da Tarski, $\mathbf{B}/F$ coinciderà con l'algebra semplice **2** *esattamente* nel caso F sia un ultrafiltro cosicché gli ultrafiltri si possono caratterizzare come gli antinuclei di omomorfismi su **2**.

A questo punto il rapporto tra le tecniche algebrico-topologiche di Stone e quelle puramente algebriche di Birkhoff nella costruzione di rappresentazioni per le algebre di Boole risulta evidente. Data la rappresentazione di Stone, possiamo ottenere un monomorfismo

$$\overline{\rho} : \mathbf{B} \rightarrow \mathbf{2}^I$$

la cui immagine sia un prodotto sottodiretto, semplicemente indiciando con I la famiglia degli ultrafiltri in B in modo che essa coincida con $\{F_i\}_{i \in I}$. Assegnare ad ogni $a \in \mathbf{B}$ l'insieme degli ultrafiltri che lo contengono, una volta che si sostituiscano agli insiemi le funzioni caratteristiche, ci porta ad un monomorfismo $\overline{\rho} : \mathbf{B} \rightarrow \mathbf{2}^I$ che è proprio quello richiesto. All'inverso, se partiamo dalla rappresentazione di Birkhoff per ottenere quella di Stone, basterà comporre il monomorfismo $\mu : \mathbf{B} \rightarrow \mathbf{2}^I$ con le proiezioni $\pi_i : \mathbf{2}^I \rightarrow \mathbf{2}$ per ottenere una famiglia di omomorfismi $\mu_i : \mathbf{B} \rightarrow \mathbf{2}$. Come osservato sopra, gli antinuclei di questi omomorfismi saranno filtri, ed in particolare ultrafiltri in quanto, essendo l'immagine di μ sottodiretta, ogni μ_i è suriettivo. A questo punto, considerato che μ assegna ad ogni $a \in \mathbf{B}$ un sottoinsieme di I (se passiamo dalle funzioni caratteristiche agli insiemi) avremo che ad ogni $a \in \mathbf{B}$ è associata una famiglia di ultrafiltri che lo contengono e questa non è altro che la rappresentazione di Stone.

Sarà questo stretto intergioco tra filtri, omomorfismi ed algebre quozienti a costituire almeno all'inizio il paradigma fondamentale di quel progetto di algebrizzazione della logica che a partire dalla metà degli anni trenta si andò formando, soprattutto in Polonia e di cui ci occuperemo ora.

7.3. *Metamatematica e logica algebrica*

Protagonisti di questa svolta nei rapporti fra logica e matematica astratta furono i logici della scuola di Leopoli e Varsavia, in particolare Tarski – di cui abbiamo già parlato e ampiamente riparleremo in seguito – Lindenbaum, Mordechai Waisberg (1905-1941), Stanislaw Jaskowski, Boleslaw Sobocinski e altri, non pochi dei quali perirono per mano dei nazisti nella distruzione del Ghetto di Varsavia. Va subito detto che ciò che contraddistingue

la logica algebrica che così si andò configurando dall'algebra della logica ottocentesca, prima ancora della maggiore articolazione dei metodi e delle tecniche è l'obiettivo di fondo. Scopo dei logici della scuola di Leopoli e Varsavia non era quello di utilizzare la teoria degli insiemi, la topologia, l'algebra per dare finalmente una veste esatta alla logica – come invece era stato per Boole, per Schröder e in parte per Peirce. Agli inizi degli anni trenta tanto la logica proposizionale che quella elementare avevano ricevuto sistemazioni pressoché definitive sul piano sintattico lungo le linee che l'*Ideografia* di Frege aveva segnato, e le indagini di Hilbert avevano già prospettato il progetto di un'analisi metamatematica di questi sistemi. Scopo della metamatematica di Hilbert – come avremo occasione di vedere più ampiamente in seguito nel capitolo IV – era di studiare la struttura delle varie teorie formali (e quindi anche dei diversi sistemi di logica pura) sulla base di criteri e di nozioni di carattere rigidamente *finitista* così da rispondere alle esigenze di un preciso programma fondazionale, quello formalista. Questo tanto per i sistemi classici che per quelli intuizionisti, la cui formalizzazione era stata iniziata da Heyting nel 1930.

È in esplicito riferimento a questo progetto che si situano le ricerche della scuola polacca, riprendendo però anche gli spunti di quella concezione metateorica della logica che già i postulazionisti americani e Sheffer in particolare avevano tentato. La formulazione logicista di sistemi logici e teorie all'interno di linguaggi formali è data per scontata e accettata; ciò che diviene l'obiettivo centrale è lo studio della *struttura matematica* di questi sistemi, l'indagine delle loro proprietà all'interno di una prospettiva che non pone pregiudiziali sul piano metodologico e non impone limitazioni al tipo di metodi usati. Teorie, calcoli, sistemi logici sono oggetti che hanno una struttura matematica che si può indagare con i metodi della teoria degli insiemi, della topologia, dell'algebra, come succede per le figure geometriche, gli oggetti fisici, ecc. Il parallelo con le figure geometriche è di Tarski che già alla fine degli anni venti andò sviluppando un progetto di *metamatematica* o *metodologia generale delle scienze deduttive*, che – e qui il collegamento con i postulazionisti americani, e in particolare con Sheffer, è evidente – si pone di fronte ai sistemi deduttivi con un atteggiamento radicalmente antitetico a quello hilbertiano. Per Tarski, i sistemi deduttivi non sono sistemi formali – insiemi di stringhe finite di

segni generati mediante regole effettive – ma insiemi di oggetti arbitrari isolati all'interno di una famiglia la cui cardinalità è numerabile e su cui è definita una *relazione di conseguenza*. Più precisamente, dato un insieme numerabile S (gli enunciati) e un'operazione $\mathrm{Cn} : \mathscr{P}(S) \rightarrow \mathscr{P}(S)$ (l'operatore di conseguenza) che soddisfa le condizioni:

A1 Se $X \subseteq S$, allora $X \subseteq \mathrm{Cn}\,(X)$

A2 Se $X \subseteq S$, allora $\mathrm{Cn}(\mathrm{Cn}(X)) = \mathrm{Cn}\,(X)$

A3 Se $X \subseteq S$, allora $\mathrm{Cn}(X) = \bigcup_{Y \in \mathscr{P}_\omega(X)} \mathrm{Cn}\,(Y)$

dove $\mathscr{P}_\omega(X)$ indica la collezione delle parti *finite* di X, un *sistema deduttivo* è ogni $X \subseteq S$ per cui $X = \mathrm{Cn}(X)$, cioè ogni X *chiuso* rispetto al passaggio alla conseguenza. In questa prospettiva non conta il carattere concreto degli enunciati, il loro essere successioni finite di simboli, né il modo in cui la relazione Cn, che ad ogni X associa le sue conseguenze, è definita, gli assiomi o le regole di inferenza che si assumono. Come un gruppo può essere presentato in tanti modi diversi, scegliendo diversi insiemi di generatori e diverse relazioni definitorie, ma ha proprietà che sono *indipendenti* dalle diverse presentazioni, così esistono proprietà dei sistemi deduttivi (come la coerenza, la completezza, ecc.) che non dipendono dalla scelta di assiomi e regole. Sono queste proprietà che la metamatematica studia con metodi che non debbono in alcun modo essere sottoposti a restrizioni di tipo finitista così come succede nella topologia, nell'algebra e nella teoria degli insiemi. Non deve stupire quindi – come avremo occasione di vedere – che Tarski come tutti i logici di cui ci occuperemo faccia all'occorrenza ricorso all'assioma di scelta di Zermelo ed ad altre ipotesi insiemistiche forti. Anzi, sarà proprio in questo periodo che in Polonia insiemisti come Sierpinski e topologi come Kuratowski cominceranno ad indagare quanto dell'Analisi e della topologia su base insiemistica dipenda da assunzioni come l'assioma di scelta e l'ipotesi del continuo, e come queste assunzioni si possano formulare in modi diversi in contesti specifici ed eventualmente indebolire.

È ad esempio Kuratowski che nel 1922 – seguendo precedenti indicazioni di Felix Hausdorff – formula un principio equivalente all'assioma di scelta che ha il grande vantaggio di potersi applicare in contesti matematici determinati senza dover far riferimento

(come – si ricorderà – faceva Stone nella sua dimostrazione del teorema dell'ultrafiltro) ad indiciazioni in termini di numeri ordinali. Il *principio di Kuratowski* afferma che in un insieme ordinato, ogni *catena* (vale a dire ogni sottoinsieme *linearmente* ordinato) è sottocatena di una catena massimale. Lungo la stessa linea, nel 1935 Max Zorn formulerà il lemma che porta il suo nome, secondo il quale se in un insieme ordinato ogni catena ha un maggiorante, allora l'insieme contiene elementi massimali. Nella stessa prospettiva, ma partendo da concetti differenti, si muoveranno nello stesso giro d'anni Oswald Teichmüller (1913-43) e J.W. Tukey che daranno altre forme equivalenti dell'assioma di scelta più direttamente adeguate a contesti algebrici o topologici.

Una volta concepiti i sistemi deduttivi nei termini astratti sopra descritti, non deve stupire quindi che, proprio come nell'algebra e nella topologia, i metodi insiemistici si rivelassero essenziali. Come ricordato, i primi risultati della metamatematica tarskiana riprendono la problematica iniziata dai postulazionisti americani affrontando in generale questioni di completezza, indipendenza, definibilità, ecc. per i sistemi deduttivi. Uno dei primi risultati, provato da Lindenbaum, è che ogni sistema deduttivo X *coerente* (tale cioè che $Cn(X) \neq S$) ammette un'estensione *massimale* nel senso che non ha a sua volta un'estensione propria che sia ancora coerente. Il teorema di Lindenbaum, come il teorema dell'ultrafiltro con cui ha stretti rapporti, è dimostrato ricorrendo all'assioma di scelta e costituisce un esempio illuminante del carattere infinitario della metamatematica di indirizzo tarskiano.

Lo stesso atteggiamento astratto lo ritroviamo d'altra parte anche in altri contesti: sia che si indaghino diverse forme di assiomatizzabilità, come di indipendenza o di definibilità, l'attenzione è sempre rivolta agli aspetti più generali dei sistemi deduttivi così che i risultati non solo non dipendano da particolari formulazioni, ma siano anche indipendenti dal tipo di linguaggio (scelta dei connettivi ed in generale del vocabolario logico ed extralogico) e dalla natura stessa delle regole logiche assunte. Non dimentichiamo infatti, come vedremo meglio in seguito nel capitolo IV, che fondatore della scuola di Varsavia fu quello stesso Jan Łukasiewicz che nel 1920 aveva dato inizio alle ricerche sulle logiche *polivalenti* (a tre valori prima e poi ad n e ad infiniti valori). Ai primi degli anni trenta risale d'altra parte la formalizzazione della logi-

ca intuizionistica data da Heyting, mentre già dai primi anni venti avevano fatto la loro comparsa negli Stati Uniti le logiche *modali* di C. I. Lewis ed i *calcoli proposizionali a più valori* di Emil Leon Post (1897-1954).

Ritorneremo più avanti su questi sviluppi; quello che vogliamo sottolineare è il ruolo che l'esistenza di sistemi logici diversi ha avuto nel modellare l'approccio astratto alla metamatematica di Tarski. La ricerca di definizioni generali per i concetti metalogici fondamentali come pure il desiderio di trovare un contesto comune in cui inquadrare le diverse possibilità e vederne i rapporti costituì un naturale incentivo alla ricerca di strutture comuni sottogiacenti ai sistemi deduttivi.

In questa ricerca, l'incontro con i nuovi metodi dell'algebra e della topologia era inevitabile e risulta evidente già dalla stessa definizione di sistema deduttivo data da Tarski. Operatori che oltre alle condizioni di A1 e A2 della definizione di operatore di conseguenza soddisfano la condizione

AM $\qquad$ Se $X \subseteq Y$ allora $\mathrm{Cn}(X) \subseteq \mathrm{Cn}(Y)$

(che segue da A3 e che Tarski non postulò esplicitamente) sono noti oggi come *operatori di chiusura* e, come mostrato da Ore alla metà degli anni quaranta, permettono da una parte di costruire una teoria generale delle corrispondenze di Galois che generalizza quella per i campi, dall'altra sono alla base di tutta la trattazione unitaria del concetto astratto di indipendenza, centrale nell'analisi dei reticoli gemetrici e nella costruzione algebrica della geometria proiettiva.

Ma c'è dell'altro. Già nel 1922 Kuratowski aveva mostrato come, dato un insieme X, se si definiva un'operazione $\mathrm{C}: \mathscr{P}(X) \rightarrow \mathscr{P}(X)$ che oltre alle condizioni A1, A2 e AM soddisfaceva:

AD $\qquad$ $\mathrm{C}(Y \cup Z) = \mathrm{C}Y \cup \mathrm{C}Z$

AV $\qquad$ $\mathrm{C}\,\emptyset = \emptyset$

si otteneva una topologia di *chiusi* considerando come tali gli insiemi $Y \in \mathscr{P}(X)$ per cui $\mathrm{C}Y = Y$. Le proprietà di C così isolate corrispondono infatti a quelle valide in uno spazio topologico, una volta che si consideri l'operatore che ad ogni insieme Y associa la sua *chiusura topologica*, l'intersezione di tutti i chiusi che lo estendono.

Ciò che emergeva in questo modo era che gli operatori C di sopra (da allora noti come *operatori di Kuratowski*) non solo erano operatori di chiusura nel senso di Ore, ma fornivano un modo unitario, alternativo a quello usuale, di definire gli spazi topologici.

Il vantaggio di questa impostazione – come Kuratowski poneva in luce – era che si apriva la strada ad uno studio su base algebrica della topologia generale. Era uno degli obiettivi per cui era nata la teoria dei reticoli e non stupisce che, una volta emigrato negli USA, sia stato proprio Tarski in collaborazione con McKinsey a sviluppare nel 1944 questo progetto di un'algebra della topologia in termini di particolari strutture, dette appunto *algebre di chiusura*, che sono individuate da un'algebra di Boole e da un operatore che soddisfa A1, A2, AM. Vedremo più avanti come queste algebre abbiano fornito lo sfondo algebrico per analizzare la logica modale; per ora basti osservare che in questo modo il concetto di sistema deduttivo veniva a porsi all'interno di una teoria generale che vedeva come casi particolari anche gli spazi topologici e si differenziava per il suo carattere di finitezza codificato dalla clausola A3 di sopra, che riflette l'idea che se $\mathscr{A}$ è conseguenza di X, per dimostrarlo si fa uso di una parte finita di assiomi in X. Nel 1935 Tarski dedicherà ben due lavori allo sviluppo del calcolo dei sistemi ed i suoi risultati, se da una parte inauguravano la classificazione delle teorie in termini di modelli, dall'altra fornivano contributi alla teoria dei reticoli di chiusi rispetto a operatori di Ore, validi indipendentemente dalle applicazioni logiche.

Il terreno principale cui la scuola di Leopoli e Varsavia applicò la sua attenzione è costituito però dallo studio della *logica proposizionale* ed in particolare dalla edificazione di una teoria generale su base algebrica e topologica il cui obiettivo era la classificazione dei diversi sistemi logici. Il primo resoconto sullo stato delle ricerche fu pubblicato nel 1930 con il titolo *Untersuchungen über den Aussagenkalkül* (*Ricerche sul calcolo proposizionale*): firmato da Łukasiewicz e Tarski, esso raccoglieva lavori di tutto il gruppo, in particolare contributi ottenuti in collaborazione con Lindenbaum. Qui, oltre a specifici risultati sulle logiche polivalenti su cui torneremo nel capitolo IV, comparivano per la prima volta due concetti che avrebbero avuto un ruolo fondamentale nel collegare i calcoli proposizionali alla teoria dei reticoli e quindi nel generale processo di algebrizzazione della logica.

Il primo è quello di *matrice logica*, la cui origine, come sappiamo, risale a Peirce e Schröder e che non solo era stato lo strumento di base – alternativo all'analisi sintattica – nello studio delle proprietà metamatematiche dei calcoli (indipendenza, completezza funzionale del sistema dei connettivi, ecc.) ma anche il punto di partenza tanto per la definizione delle logiche polivalenti di Łukasiewicz quanto per quelle di Post. È appunto questo secondo aspetto delle matrici, quello cioè di metodo per generare e individuare sistemi logici, che Tarski poneva in primo piano. Per definizione, una *matrice* **M** per un linguaggio proposizionale L in cui compaiono i soli connettivi $C_1, \ldots, C_k$ è data da una struttura

$$< A, D, \{\overline{C}_i\}_{i \leq k} >$$

dove A è un insieme (la famiglia dei valori possibili) $D \subseteq A$ un sottoinsieme (che individua i valori *designati*) e, per ogni $i \leq k$, $\overline{C}_i : A^n \rightarrow A$ è una funzione a n argomenti (se n sono gli argomenti del connettivo C_i).

In questa forma le matrici costituiscono una generalizzazione dei *sistemi deduttivi* come erano stati concepiti da Huntington e Bernstein in quanto si possono vedere come individuate da un'*algebra* $\mathbf{A} = <A, \{\overline{C}_i\}_{i \leq k}>$ ed un sottoinsieme D di A e basta che $\mathbf{A}$ sia un'algebra di Boole e D un filtro su $\mathbf{A}$ per ottenere i sistemi che Bernstein aveva in mente. In questa ottica, prototipo di filtro in un'algebra diverrà la collezione degli enunciati veri in una data teoria, di ideale quella degli enunciati falsi. Il fatto importante è che in questo modo abbiamo da una parte un linguaggio proposizionale L le cui proposizioni atomiche possiamo assumere siano date da un insieme $\{p_i\}_{i \in I}$, dall'altra un'algebra $\mathbf{A}$ le cui operazioni sono in corrispondenza biunivoca con i connettivi e con un sottoinsieme D costituito dai valori designati. Su questa base è possibile edificare una analisi *semantica* di L. Interpretare L in $\mathbf{A}$ significherà infatti dare una funzione $v : \{p_i\}_{i \in I} \rightarrow \mathbf{A}$ (in termini tecnici, una *valutazione*) ed estenderla ad una interpretazione di *tutte* le formule di L ponendo per definizione che, per ogni $i \leq k$:

$$\overline{v}(C_i(a_i, \ldots, a_n)) = \overline{C}_i(\overline{v}(a_1), \ldots, \overline{v}(a_n)).$$

In questo modo – come già avevano fatto Peirce e Schröder – ai

connettivi C_i viene associata nell'interpretazione una *funzione di verità* $\overline{C}_i$. Si tratta ora di definire quando una formula a risulta *vera* nella matrice **M** rispetto alla valutazione v. È a questo livello che interviene l'insieme dei valori designati, e *vera* in **M** rispetto a v sarà ogni formula $\mathscr{A}$ tale che $\bar{v}(\mathscr{A})$ appartiene a D. **M**-*tautologia* sarà ogni formula $\mathscr{A}$ che risulta vera in *ogni* valutazione delle formule atomiche. Viene così associato naturalmente ad ogni matrice **M** un insieme $E(\mathbf{M})$ di formule costituito da tutte e sole le proposizioni che risultano tautologie rispetto alla matrice data. Stesso discorso si può fare, generalizzando, ponendo al posto di una singola matrice **M** una famiglia Γ di matrici. Quel che conta è che per ampie classi di famiglie Γ di matrici si può vedere che, se definiamo che $\mathscr{A} \in \mathrm{Cn}_\Gamma(X)$ (in parole: che $\mathscr{A}$ è una *Γ–conseguenza tautologica* di X) se $\mathscr{A}$ assume valore designato in ogni valutazione rispetto alla quale assumono valore designato tutte le formule in X, Cn_Γ soddisferà le condizioni che definiscono gli operatori di conseguenza, individuando così sistemi deduttivi nel senso di Tarski.

In questo modo si veniva a disporre di un metodo puramente algebrico per costruire sistemi deduttivi proposizionali (e quindi sistemi logici) che si contrapponeva a quello puramente sintattico di origine logicista fondato su assiomi e regole. Nuovi calcoli e nuove logiche si potevano individuare variando le algebre e gli insiemi di valori designati. Di fatto Łukasiewicz e Tarski si limitavano a linguaggi con i soli connettivi $\neg$ e $\rightarrow$ e consideravano operazioni di conseguenza definite in termini di specifiche operazioni su formule: il *modus ponens* (per cui se $\mathscr{A} \in X$ e $(\mathscr{A} \rightarrow \mathscr{B}) \in X$ allora anche $\mathscr{B} \in X$) e la *sostituzione* (per cui se $\mathscr{A} \in X$ allora anche ogni formula $\mathscr{A}'$ che si ottiene da $\mathscr{A}$ rimpiazzando uniformemente sottoformule atomiche con formule arbitrarie apparterrà ad X). L'insieme delle conseguenze $\mathrm{Cn}(X)$ di un insieme X era definito come l'intersezione di tutti gli insiemi chiusi rispetto alle due operazioni e che includono l'insieme X. Se la limitazione al modus ponens era occasionale, il riferimento alla sostituzione aveva il ruolo di sottolineare la natura *logica* delle relazioni di conseguenza considerate, la loro indipendenza dal significato delle singole formule atomiche. In questo senso i sistemi ottenuti si potevano considerare sistemi *logici* e la domanda naturale diventava quella di sapere entro che limiti il metodo delle matrici era *universale*, permetteva cioè di ottenere *tutti* i possibili sistemi logici.

La risposta venne data da Lindenbaum introducendo quelle che oggi sono note come *algebre di Lindenbaum* e che per la prima volta furono descritte esplicitamente in un lavoro congiunto da McKinsey e Tarski nel 1948. È questo il secondo concetto fondamentale cui facevamo riferimento più sopra e che permette la completa traduzione in termini algebrici dello studio dei sistemi logici proposizionali. Per valutarne appieno la portata consideriamo un linguaggio proposizionale L con formule atomiche $\{p_i\}_{i \in I}$ e i connettivi $\wedge, \vee, \neg, \rightarrow$. L'idea è di vederlo come una struttura algebrica – la cosiddetta *algebra delle formule* – prendendo come dominio l'insieme S di tutte le formule e definendo per ogni connettivo una operazione corrispondente. Nel nostro caso avremo così:

$$\wedge(\mathscr{A}, \mathscr{B}) = (\mathscr{A} \wedge \mathscr{B}); \vee(\mathscr{A}, \mathscr{B}) = (\mathscr{A} \vee \mathscr{B});$$
$$\rightarrow(\mathscr{A}, \mathscr{B}) = (\mathscr{A} \rightarrow \mathscr{B}); \neg(\mathscr{A}) = (\neg \mathscr{A}).$$

Come si vede, l'algebra non fa altro che mettere in luce il meccanismo con cui le formule complesse sono ottenute da formule più semplici. Supponiamo ora di avere un sistema *logico* (nel senso sopra indicato) $C(X)$. Possiamo allora definire tra le formule del dominio la relazione

$$\mathscr{A} \equiv \mathscr{B} \quad \text{se e solo se} \quad (\mathscr{A} \rightarrow \mathscr{B}), (\mathscr{B} \rightarrow \mathscr{A}) \in C(X).$$

Sotto ipotesi estremamente generali riguardanti il sistema deduttivo $C(X)$, si può provare non solo che $\equiv$ è una relazione di equivalenza ma addirittura che è una congruenza rispetto alle operazioni dell'algebra. La cosa non è strana se si pensa che, intuitivamente, la $\equiv$ corrisponde alla *equivalenza logica* rispetto al sistema, cosicché avremo, esprimendoci direttamente in termini di formule, che:

Se $\mathscr{A} \equiv \mathscr{A}'$ e $\mathscr{B} \equiv \mathscr{B}'$ allora

$$(\mathscr{A} \wedge \mathscr{B}) \equiv (\mathscr{A}' \wedge \mathscr{B}'), (\mathscr{A} \vee \mathscr{B}) \equiv (\mathscr{A}' \vee \mathscr{B}')$$
$$(\mathscr{A} \rightarrow \mathscr{B}) \equiv (\mathscr{A}' \rightarrow \mathscr{B}'), (\neg \mathscr{A}) \equiv (\neg \mathscr{A}').$$

Possiamo allora quozientare l'algebra delle formule Form_L ed ottenere l'algebra $\mathbf{A}_{C(X)} = \mathrm{Form}_L / \equiv$. È questa l'*algebra di Lindenbaum* associata al sistema logico e a questo punto è facile verificare

che, se con $\{X/\equiv\}$ indichiamo l'insieme delle classi di equivalenza delle formule in X, la matrice $\mathbf{M}_{C(X)} = <\mathbf{A}_{C(X)}, \{X/\equiv\}>$ sarà una matrice *caratteristica* per il sistema, nel senso che avremo:

$$\mathscr{A} \in C(X) \text{ se e solo se } \mathscr{A} \in E(\mathbf{M}_{C(X)}).$$

È questo il risultato di Lindenbaum e nella sua estrema semplicità esso ha conseguenze di notevole importanza.

I sistemi deduttivi, come sappiamo, possono essere definiti in modi estremamente vari ed in particolare in termini *sintattici*, come calcoli (secondo la linea fregeana) dati da assiomi e da regole di inferenza. Nella scelta di assiomi e regole compaiono ovviamente considerazioni che hanno a che fare con nozioni intuitive di verità, interpretazione, ecc., ma queste considerazioni *semantiche* non sono oggetto esplicito dell'analisi e della costruzione sintattica dei sistemi. Matrici, valutazioni ed algebre di Lindenbaum permettono invece di dare forma ad una vera e propria *semantica* per i linguaggi proposizionali ed in particolare di stabilire un collegamento tra calcoli e nozioni semantiche di verità. Come scrivevano Łukasiewicz e Tarski nel loro rapporto «entrambi i metodi hanno i loro vantaggi e svantaggi. I sistemi costruiti con il metodo assiomatico [come calcoli] sono più facili da analizzare per quanto riguarda la loro assiomatizzabilità, ma i sistemi generati da matrici sono più facili da studiare quando si considerano coerenza e completezza». Il significato fondamentale del teorema di Lindenbaum è che esso mostra come, sotto ipotesi molto generali, *tutti* i calcoli si possono analizzare in termini di matrici e quindi ne è possibile uno studio semantico. Ciò da una parte rende più intuitivo (in quanto fa riferimento a nozioni di verità) il contenuto logico dei sistemi, dall'altra permette un più diretto studio di certe proprietà metamatematiche.

Centrale in questa prospettiva diviene allora il problema di trovare insiemi di matrici Γ che siano *caratteristiche* per un dato sistema $C(X)$ così che si abbia $C(X) = E(\Gamma)$ e le conseguenze coincidano con le Γ-tautologie. Fatto questo, il problema successivo è di vedere se ci si può ridurre ad una matrice sola e se questa è particolarmente dominabile dal punto di vista matematico, in particolare se è *finita*, cosicché – come nel caso indagato da Peirce e Schröder – il metodo delle tavole di verità ci dà un mo-

do per decidere quando una formula è o meno legge del sistema. Veniva così naturale introdurre operazioni su matrici e studiare l'effetto che esse hanno sull'insieme delle tautologie; queste operazioni a loro volta rimandavano ad operazioni sulle algebre soggiacenti e ci si trovava così direttamente nel contesto dell'algebra universale e – per buona parte delle logiche considerate – della teoria dei reticoli. La teoria generale delle matrici avrebbe trovato una sistemazione definitiva nell'opera di Jerzy Łos nel 1949, quando si sarebbe sviluppata una seconda fase della metodologia delle scienze deduttive; nel periodo di mezzo il processo di algebrizzazione si sarebbe raffinato stabilendo un completo *vocabolario di traduzione* tra concetti e proprietà metalogiche e proprietà algebriche.

Per illustrare meglio questo fatto consideriamo il caso della logica classica. Supponiamo allora che **K** sia il sistema logico proposizionale classico in una sua specifica assiomatizzazione, poniamo, quella data da Hilbert e Ackermann nel 1928 e – come oggi è consueto – scriviamo $\vdash \mathscr{A}$ per indicare che $\mathscr{A}$ è un teorema logico, cioè che $\mathscr{A} \in \mathrm{Cn}(X)$, dove X è l'insieme degli assiomi del sistema. Applicando il procedimento illustrato sopra, da **K** otterremo l'algebra di Lindenbaum $\mathbf{A}_{\mathbf{K}}$. Non è una sorpresa constatare che essa è un'algebra di Boole e che apparterrà così alla varietà $V_{\mathbf{B}}$ delle algebre che soddisfano gli assiomi che abbiamo visto a suo tempo. Meno ovvio è che essa risulta *libera* in questa classe sulla famiglia $X = \{[p_i]\}_{i \in I}$ di *generatori* data dalle classi di equivalenza delle formule atomiche.

Il concetto di algebra libera è uno degli strumenti centrali dell'algebra universale ed era stato introdotto da Birkhoff alla metà degli anni trenta, generalizzando l'analogo concetto che compare nella teoria dei gruppi. Ancora Birkhoff aveva mostrato come in ogni varietà esistessero algebre libere su ogni insieme di generatori e come, fissata una cardinalità, in una varietà ci fosse una sola algebra libera (a meno di isomorfismi) su un insieme di generatori di quella cardinalità. La centralità delle algebre libere sta nel loro carattere di universalità che in certo senso le rende rappresentative dell'intera varietà cui appartengono, analoghi algebrici della teoria equazionale della classe. Per definizione, infatti, un'algebra **A** appartenente ad una varietà V è *libera in V sull'insieme di generatori X* se è *generata* da X (nel senso che tutti i suoi elementi si ottengono

da elementi di X applicando le operazioni dell'algebra) e per ogni funzione f: $X \rightarrow B$, dove B è il dominio di un'algebra in V, f si estende in modo unico ad un omomorfismo $\bar{f}: \mathbf{A} \rightarrow \mathbf{B}$. Questo significa che nessuna equazione $t(x_1,...,x_n) = s(x_1,...,x_n)$ tra termini che coinvolgono generatori potrà essere vera nell'algebra libera e falsa in una qualunque algebra della classe. È in questo senso che le algebre libere rappresentano la varietà ed è questo che permette di dimostrare che data una varietà ogni sua algebra sarà quoziente di un'opportuna algebra libera.

Il fatto che la nostra algebra di Lindenbaum **A** sia libera sull'insieme $\{[p_i]\}_{i \in I}$ nella varietà delle algebre di Boole, ci consente allora di compiere un passo decisivo e di identificare esplicitamente interpretazioni (valutazioni) di L in una matrice **M** con algebra di Boole sottogiacente **B**, con omomorfismi $\mu : \mathbf{A_K} \rightarrow \mathbf{B}$. Ogni valutazione v infatti ci dà una funzione $f: X \rightarrow \mathbf{A}$ definita ponendo $f([p_i]) = \bar{v}(p_i)$. Sfruttando il fatto che $\mathbf{A_K}$ è libera, avremo immediatamente l'omomorfismo $\bar{f} : \mathbf{A_K} \rightarrow \mathbf{A}$ e per definizione è facile verificare che $\bar{f}$ $([\mathscr{A}])$ coinciderà con $\bar{v}$ $(\mathscr{A})$ per ogni formula $\mathscr{A}$.

In altre parole: è possibile tradurre *completamente* ogni discorso sulle interpretazioni (con valori in matrici su algebre di Boole) in termini di omomorfismi dall'algebra di Lindenbaum ad algebre in $V_{\mathbf{B}}$. Ma c'è di più. Per definizione ogni sottoinsieme T delle formule avrà il suo corrispondente insieme quozientato $T/\equiv$ in $\mathbf{A_K}$. Ci domandiamo: che cosa corrisponde dal punto di vista algebrico ad insiemi di formule che sono significativi dal punto di vista logico? Uno dei concetti logici fondamentali è sicuramente quello di *teoria*, dove teoria (rispetto al calcolo **K**) è ogni insieme T di formule chiuso rispetto al passaggio ai teoremi dimostrati sulla base di **K**. Se $\mathrm{Cn}_{\mathbf{K}}$ è l'operazione che ad ogni T associa la famiglia delle formule che si ottengono da quelle di T applicando le regole di **K** ed usando gli assiomi logici, avremo che teoria non è altro che un sistema deduttivo $T = \mathrm{Cn}_{\mathbf{K}}(T)$, secondo la definizione generale di Tarski (in quest'ottica, l'insieme dei teoremi del calcolo coinciderà con $\mathrm{Cn}_{\mathbf{K}}$ $(\emptyset)$). È adesso facile esercizio verificare che se T è una teoria (rispetto a **K**), $T/\equiv$ sarà un filtro, cosicché saranno i filtri le controparti algebriche delle teorie. Quel che è ancora più interessante è che la classificazione dei filtri corrisponde in questo modo alla classificazione delle teorie: *coerenti* saranno tutte e sole le teorie che corrispondono a filtri propri, mentre *complete* (*sintatticamente*,

tali cioè che per ogni $\mathscr{A}$ o $\mathscr{A}$ oppure $\neg\ \mathscr{A}$ – ma non entrambe – è teorema) saranno quelle che corrispondono a filtri *massimali*, agli ultrafiltri. Più in generale avremo che se X è un insieme di formule

$$\mathscr{A} \in \mathrm{Cn}_{\mathrm{K}}(X) \qquad \text{se e solo se} \qquad [\mathscr{A}] \in [X/\equiv]_{\mathbf{A}_{\mathrm{K}}}$$

dove $[X/\equiv]_{\mathbf{A}_{\mathrm{K}}}$ è il filtro generato da $X/\equiv$ in $\mathbf{A}_{\mathrm{K}}$.

A questo punto il dizionario di traduzione è completo e non riguarda solo il calcolo classico e le algebre di Boole, ma ragionamenti analoghi si possono applicare alle logiche modali, a quella intuizionista, alla quasi totalità dei calcoli logici proposizionali.

L'utilità di un simile vocabolario di traduzione non tardò ad emergere una volta che si affrontarono espliciti problemi metateorici sui calcoli logici, primo fra tutti quello della *completezza* (*semantica*, detta talora anche *adeguatezza*, anche se quest'ultimo termine non è mai divenuto di uso generale). Per la logica classica proposizionale un problema del genere era già stato posto da Hilbert e Ackermann nel 1928: entro che limiti possiamo affermare che i teoremi del calcolo *coincidono* con le tautologie a due valori? È immediato che tutti i teoremi sono tautologie, cioè che $\mathrm{Cn}(\emptyset) \subseteq E\,(\mathbf{M}_2)$, dove $\mathbf{M}_2 = <\mathbf{2}, \{1\}>$, ma più complesso è provare l'inverso, vale a dire il *teorema di completezza* per cui $\mathbf{M}_2$ è una matrice *caratteristica* per il calcolo. Già Post – come vedremo più ampiamente nel capitolo IV – aveva risolto il problema utilizzando le forme normali, ma è interessante vedere come questo si possa fare in termini algebrici utilizzando filtri e teoremi di rappresentazione; sarà questa la strada che a partire dalla metà degli anni trenta sarà seguita per mostrare la completezza dei calcoli logici non classici rispetto a date interpretazioni semantiche.

Le vie aperte sono sostanzialmente due e passano attraverso le algebre di Lindenbaum. Procediamo per assurdo e supponiamo che $\nvdash_{\mathrm{K}} \mathscr{A}$; per provare la completezza del calcolo rispetto alla sola matrice $\mathbf{M}_2$ basterà provare che $\mathscr{A}$ non è $\mathbf{M}_2$-tautologia. Consideriamo allora l'algebra di Lindenbaum $\mathbf{A}_{\mathrm{K}}$. Avremo che la $[\mathscr{A}] \neq 1$ in quanto, per il collegamento visto sopra tra teorie e filtri, $[\mathscr{A}]$ non potrà appartenere al filtro generato da 1. Ricorriamo ora al teorema di rappresentazione di Birkhoff. Avremo allora un monomorfismo $\mu : \mathbf{A}_{\mathrm{K}} \to \mathbf{2}^I$ che manderà $[\mathscr{A}]$ su un elemento diverso da 1. Componiamo quindi con un'opportuna proiezione $\pi_i : \mathbf{2}^I \to \mathbf{2}$.

L'omomorfismo composto corrisponderà ad una valutazione v e poiché assegnerà ad $[\mathscr{A}]$ un valore diverso da 1, avremo che nella valutazione v, $\bar{v}(\mathscr{A}) \neq 1$ e quindi $\mathscr{A}$ non è $\mathbf{M_2}$-tautologia. Non diverso sarebbe stato il ragionamento se fossimo partiti dalla rappresentazione di Stone ma – una volta deciso di utilizzare gli ultrafiltri – esiste una strada più diretta che fu seguita per la prima volta da Łos nel 1949. Supposto di avere un'algebra di Lindenbaum $\mathbf{A_K}$ e un elemento $[\mathscr{A}] \neq 1$, sappiamo per il teorema dell'ultrafiltro di Stone che esisterà un filtro massimale F tale che $[\mathscr{A}] \in F$. Ricorriamo ora al collegamento tra ultrafiltri e omomorfismi su $\mathbf{2}$ e, dato F, possiamo definire un omomorfismo $\mu : \mathbf{A_K} \to \mathbf{2}$ il cui antinucleo sarà F, per cui $\mu\,([\mathscr{A}]) = 1$ e $\mathscr{A}$ non sarà una $\mathbf{M_2}$-tautologia.

Il fatto decisivo è che questi metodi sono molto generali e non si limitano alla logica classica. Il primo esempio d'applicazione a logiche diverse da **K** fu ancora una volta dato da Tarski nel 1938 in un famoso articolo, *Der Aussagenkalkül un die Topologie* (*Il calcolo proposizionale e la topologia*) in cui non solo individuava una semantica algebrica adeguata per il calcolo proposizionale intuizionista, ma la applicava allo studio dei rapporti tra questa e la logica classica. Il punto di partenza di Tarski era costituito dalle strutture algebriche che emergono una volta che si considerino gli *aperti* di uno spazio topologico e non semplicemente le parti qualsiasi. Dato lo spazio $\mathbf{T} = \langle X, \mathscr{O}_\mathbf{T} \rangle$, dove $\mathscr{O}_\mathbf{T}$ è la famiglia degli aperti, è facile verificare non solo che essa costituisce un reticolo distributivo limitato (massimo X, minimo $\emptyset$) ma che è possibile definire tra aperti un'altra operazione, lo *pseudocomplemento relativo* $\Rightarrow$ *che ad ogni coppia di aperti* X, Y associa il *massimo* insieme $(Y \Rightarrow Z)$ per cui $Y \cap (Y \Rightarrow Z) \subseteq Z$. L'esistenza di questo massimo è ovvia ove si pensi che – essendo aperti gli insiemi su cui operiamo – avremo che $(Y \Rightarrow Z) = \bigcup \{ W \mid Y \cap W \subseteq Z\}$.

Quello che interessa è il significato logico di questa operazione che enuclea una delle proprietà centrali del connettivo di condizionale, e cioè il fatto che $(\mathscr{A} \Rightarrow \mathscr{B})$ è la proposizione più debole che congiunta ad $\mathscr{A}$ ci dà come conseguenza $\mathscr{B}$. È questa una combinazione che riflette tanto la validità del modus ponens quanto quella di quel *principio* (o teorema) *di deduzione* che proprio in quegli anni Tarski aveva dimostrato per la logica classica, ma che vale anche per la logica intuizionista. Utilizzando un

termine che sarebbe stato introdotto più tardi, possiamo allora dire che la famiglia degli aperti di uno spazio topologico costituisce un'*algebra di Heyting* e questo fatto, con il conseguente collegamento fra logica intuizionista e topologia, era stato notato indipendentemente da Stone nel 1937 in un articolo dedicato alle rappresentazioni topologiche dei reticoli distributivi.

Per vedere meglio di che collegamento si tratti, osserviamo che in ogni algebra di Heyting $\mathbf{H}$ è possibile definire dallo pseudocomplemento relativo uno pseudocomplemento *assoluto* $\dot{-}$, *ponendo* $\dot{-}\, Y = (Y \Rightarrow 0)$. È immediato verificare che, per definizione, $Y \wedge \dot{-}\, Y = 0$, mentre in generale non varrà l'analogo del terzo escluso $Y \vee \dot{-}\, Y = 1$. Ciò risulta chiaro una volta che – prendendo come esempi proprio algebre di aperti – si vede che in questo caso avremo

$$\dot{-}\, Y = \mathrm{J}(\sim Y)$$

dove J è l'operatore che ad ogni insieme Z associa il suo *interno*, ossia la riunione degli aperti W inclusi in Z. In generale $\mathrm{J}(Y) \subseteq Y$, così che, tranne nel caso J si riduca all'identità, $\dot{-}\, Y$ non potrà coincidere con l'intero complemento insiemistico. Come mostrava Tarski, *tutti* i teoremi intuizionisti risultavano tautologie rispetto alle matrici la cui algebra sottogiacente era un'algebra di aperti e l'insieme degli elementi designati si riduceva all'unità. I teoremi classici si potevano caratterizzare come le tautologie rispetto a quegli spazi degeneri in cui l'operatore di interno coincide con l'identità (i cosiddetti spazi *isolati*) in quanto le algebre di Boole non sono altro che algebre di Heyting in cui vale $Y \vee \dot{-}\, Y = 1$ e d'altra parte, come risulta dal fatto che l'algebra di Lindenbaum $\mathbf{A_K}$ per il calcolo classico è libera in $V_{\mathbf{B}}$, i teoremi classici sono tautologie in ogni matrice la cui algebra sottogiacente è di Boole e $D = \{1\}$.

Tarski riusciva a provare anche il converso del risultato di sopra – e cioè che esistono spazi per cui i teoremi del calcolo intuizionista *coincidono* con le tautologie rispetto all'algebra degli aperti associata – sfruttando alcuni risultati di Jaskowski ottenuti nel 1936. Jaskowski aveva definito una *famiglia* $\{\mathbf{I}_n\}_{n \in \mathbf{N}}$ di matrici finite che risultava caratteristica per il calcolo intuizionista. Ciò che Tarski fece vedere è che queste matrici si potevano ottenere in termini topologici e che esistevano spazi in cui tutte le loro algebre erano rappresentabili come particolari sottoalgebre di aperti. Co-

me Gödel aveva dimostrato (nel 1933) non era possibile sperare in meglio, in quanto il calcolo intuizionista non poteva avere matrice caratteristica *finita*.

Con questo rimaneva aperto il problema di fornire un *metodo di decisione* per la logica proposizionale intuizionista, in quanto la non esistenza di matrici caratteristiche finite sembrava rendere inutilizzabile il ricorso a tavole di verità. Sarà McKinsey nel 1941 a fornire una soluzione di tipo algebrico dimostrando come, per ogni formula $\mathscr{A}$ non tautologica, fosse sempre possibile trovare una matrice *finita* che la falsificava, la cui cardinalità $k_{\mathscr{A}}$ era computabile a partire dal numero delle sottoformule di $\mathscr{A}$. In questo modo, data $\mathscr{A}$, si poteva concludere che essa era teorema intuizionista se e solo se risultava verificata in tutte le matrici *finite* con algebra di cardinalità inferiore o uguale a $k_{\mathscr{A}}$. Anche se praticamente non implementabile, era questo un approccio ai problemi di decisione che avrebbe avuto grandissimi sviluppi in quanto numerosissimi calcoli godono della cosiddetta *proprietà del contromodello finito* nel senso che se una formula non è teorema si può fissare un limite massimo *finito* alla cardinalità dei possibili contromodelli.

Lo studio di contromodelli finiti è un ulteriore esempio dell'utilità delle tecniche algebriche anche al di fuori dei problemi di completezza. Esempi altrettanto significativi dava lo stesso Tarski nell'articolo citato mostrando come i teoremi classici si potessero caratterizzare all'interno di quelli intuizionisti come formule che in ogni interpretazione su algebre d'aperti avevano come valore un insieme *denso* (la cui chiusura cioè coincide con l'intero spazio). Sulla stessa linea Tarski mostrava che gli aperti regolari costituiscono in ogni algebra di aperti una sottoalgebra di Boole; risultato questo – come il precedente – che avrebbe poi ulteriormente precisato nel 1945 in collaborazione con McKinsey, dimostrando che l'algebra degli aperti regolari di uno spazio topologico è isomorfa al quoziente dell'algebra degli aperti modulo il filtro degli aperti densi.

Questo collegamento fra algebre di Heyting e di Boole non aveva solo un significato matematico, ma comportava conseguenze logiche immediate. Per definizione, infatti, un aperto regolare è un insieme Y per cui $\dot{-}\ \dot{-}\ Y = Y$ e densi saranno invece gli aperti Y per cui $\dot{-}\ Y = 0$. Questi fatti permettevano di dimostrare in modo puramente algebrico il teorema – provato da V. Glivenko nel 1929

– secondo il quale una formula $\mathscr{A}$ è teorema classico se e solo se $\neg\neg\mathscr{A}$ è teorema intuizionista. Come avremo occasione di vedere meglio nel capitolo IV, quello di Glivenko era il primo di una lunga serie di risultati che mostravano come una formula $\mathscr{A}$ sia dimostrabile classicamente se e solo se una sua opportuna traduzione $\mathscr{A}^*$ lo è intuizionisticamente. Una volta caratterizzati in termini di matrici i teoremi dei due calcoli, tutti questi risultati si potevano affrontare definendo opportune operazioni che da algebre di Heyting ci portano ad algebre di Boole e come tali non tardarono ad essere affrontati. Lo stesso – come mostrarono ancora Tarski e McKinsey nel 1948 – valeva per la *proprietà della disgiunzione del calcolo intuizionista* per cui $a \vee b$ è dimostrabile se e solo se lo è uno dei disgiunti, il che significa, in altre parole, che i teoremi del calcolo costituiscono una teoria *prima*. In questo caso scattava la possibilità di relativizzazione delle topologie.

Se i risultati che siamo andati ricordando mostrano chiaramente il configurarsi, sin dall'inizio degli anni trenta, di uno studio sistematico della logica proposizionale su base algebrica (e topologica) l'ulteriore passo decisivo verso una completa algebrizzazione della logica (proposizionale ed elementare) avvenne quando – a partire dalla metà degli anni quaranta – lo studio approfondito dei reticoli portò a teoremi di rappresentazione lungo la linea indicata da Stone e Birkhoff nel caso delle algebre di Boole.

Ancora una volta erano stati Stone e Birkhoff ad aprire la strada, il primo mostrando nel 1937 che ogni reticolo distributivo è isomorfo ad un campo d'insiemi, con una tecnica analoga a quella utilizzata per le algebre di Boole, ma sostituendo gli ultrafiltri con i filtri *primi*, il secondo provando nel 1936 (contemporaneamente e indipendentemente da Menger) che ogni reticolo modulare complementato è isomorfo al prodotto di un'algebra di Boole finita e di un numero finito di geometrie proiettive. Del 1937 erano anche alcuni risultati di H. MacNeille che – oltre a fornire un metodo generale per immergere ogni reticolo in un reticolo completo in modo da conservare tutti i supremi e gli infimi – aveva posto in luce come ogni insieme ordinato si potesse prolungare in un reticolo, mostrando quindi la generalità d'applicazione delle tecniche reticolari. A prescindere dalle applicazioni topologiche e algebriche che ebbero queste indagini, vogliamo soffermarci su quelle legate alla logica.

Abbiamo parlato delle algebre di aperti e abbiamo visto come Tarski già dal 1938 ne avesse messo in luce i legami con la logica intuizionista. Nello stesso lavoro Tarski sottolineava come anche il calcolo dei sistemi deduttivi offrisse un dominio d'interpretazione per questa logica e più in generale come questo valesse per reticoli di ideali, una volta definito opportunamente lo pseudocomplemento di un sistema X come la riunione di tutti i sistemi Y tali che $Y \cap X = \mathrm{Cn}(\emptyset)$, il sistema deduttivo minimo. Sul piano generale ciò mostrava che erano possibili algebre di Heyting definite in modo molto diverso tra loro e poneva il problema di analizzare se anche per queste algebre fosse possibile un risultato di rappresentazione.

La risposta fu data da Tarski in collaborazione con McKinsey nel 1944 e nel 1946 in due lavori che abbiamo più volte citato *The Algebra of Topology* (*L'algebra della topologia*) e *On closed elements in closure algebras* (*Sugli elementi chiusi delle algebre di chiusura*). Qui venivano definite come *algebre di chiusura* reticoli del tipo $<A, \wedge, \vee, -, C, 0,1>$ dove $< A, \wedge, \vee, -, 0,1>$ è un'algebra di Boole e C è un operatore definito su **A** che soddisfa le condizioni isolate sopra per gli operatori di Kuratowski. In questo contesto era possibile quell'analisi algebrica della topologia di cui abbiamo già più volte parlato e che era stato uno dei motivi per lo sviluppo della teoria. In termini di algebre di chiusura sono infatti esprimibili numerosi concetti centrali nella topologia senza dover ricorrere alla distinzione tra punti e zone, e ad esempio possiamo parlare di *chiusi* (gli $a \in A$ per cui $Ca = a$), di *aperti* (gli $a \in A$ per cui $- C - a = a$), di *densi*, di *regolari*, ecc. Tarski e McKinsey analizzavano in dettaglio la struttura dei chiusi e provavano che essi costituiscono un particolare tipo di reticolo, le cosiddette *algebre di Brouwer*.

Oltre a $\wedge, \vee, 0, 1$, le algebre di Brouwer sono dotate di un'operazione, la *pseudo differenza*, definita in termini dell'operatore C come $(a \Leftarrow b) = C(b \wedge -a)$ cosicché $(a \Leftarrow b)$ risulta il più piccolo elemento x dell'algebra per cui $(b \vee x) \geq a$. Per il resto, ogni algebra di Brouwer è un reticolo distributivo limitato ed è facile verificare che così definite le algebre di Brouwer non sono altro che le strutture *duali* delle algebre di Heyting, nel senso che, data un'algebra di Heyting **H**, se consideriamo la struttura **B** in cui l'ordine $\leq$ è invertito, avremo che il massimo diviene minimo, $\wedge$ diviene $\vee$ e lo pseudocomplemento relativo diviene la pseudodifferenza e vi-

ceversa. Ciò non sorprende se si considera che le algebre di Heyting nascono come algebre di aperti, che i chiusi non sono altro che complementi di aperti e che in ogni algebra di Boole $a \leq$ se e solo se $-b \leq -a$.

Tarski e McKinsey ottennero tutti i loro risultati algebrici sulla logica proposizionale intuizionista utilizzando le algebre di Brouwer e quindi considerando le valutazioni in queste algebre che *falsificano* le formule, nel senso che danno loro valore 0. Tenuto conto che – come si può verificare facilmente – l'algebra di Lindenbaum $\mathbf{A}_\mathbf{I}$ del calcolo intuizionista **I** è un'algebra di Heyting, non risulta strano che nello studio della logica intuizionista si sia ben presto passati dalle algebre di Brouwer a quelle di Heyting; solo in tempi recenti i lavori di C. Rauszer e altri hanno riportato l'attenzione sulle algebre di Brouwer in un contesto di carattere più generale.

Ciò che Tarski e McKinsey mostrarono è che ogni algebra di chiusura **C** ha una rappresentazione topologica, nel senso che esiste un monomorfismo μ che immerge **C** nell'algebra delle parti $\mathscr{B}(X)$ di uno spazio $\mathbf{T} = \langle X, \mathscr{O} \rangle$ in modo che $\mu(\mathrm{C}a) = \overline{\mu(a)}$ dove $\overline{}$ è l'operazione di chiusura che corrisponde alla topologia di **T**. Il teorema mostrava così che ogni algebra di chiusura è isomorfa a un campo d'insiemi munito di un operatore di Kuratowski cosicché ogni algebra di Brouwer risultava a sua volta isomorfa a un'algebra di chiusi di uno spazio topologico, mentre ogni algebra di Heyting lo era ad un'algebra di aperti.

Si otteneva in questo modo una nuova dimostrazione, più diretta e uniforme, della caratterizzazione precedentemente data da Tarski delle leggi logiche intuizioniste come formule valide in ogni matrice su algebre di aperti. Bastava semplicemente osservare – come già fatto sopra – che l'algebra di Lindenbaum del calcolo **I** è di Heyting. A questo punto, per provare che se una formula $\mathscr{A}$ non dimostrabile in **I** è refutabile in un'algebra di aperti, basta osservare che avremo $[\mathscr{A}] \neq 1$. Sfruttando il teorema di rappresentazione potremo trovare un isomorfismo $h : \mathbf{A}_\mathrm{I} \to \mathbf{O}$ in un'algebra di aperti così che $h([\mathscr{A}]) \neq 1$. Essendo libera su $\{p_i\}_{i \in I}$ l'algebra di Lindenbaum, come nel caso classico potremo allora concludere che abbiamo una valutazione v: $\{p_i\}_{i \in I} \to \mathbf{O}$ in un'algebra di aperti per cui $v\ (\mathscr{A}) \neq 1$: basterà porre per definizione $v\ (p_i) = h([p_i])$.

Questo schema dimostrativo (costruzione dell'algebra di Lindenbaum, dimostrazione del fatto che è libera sull'insieme delle

classi di equivalenza delle formule atomiche in una classe Γ di algebre, teorema di rappresentazione per le algebre in Γ come algebre isomorfe a quella di una sottoclasse $\Gamma_0 \subseteq \Gamma$, conclusione che i teoremi del calcolo coincidono con le formule vere in ogni valutazione in algebre Γ_0) diventerà il paradigma dell'analisi algebrica dei calcoli proposizionali e sarà utilizzato per una classe sempre più ampia di sistemi logici, primo fra tutti le logiche modali introdotte da Lewis. Negli stessi lavori che abbiamo citato Tarski e McKinsey mostravano infatti che l'algebra di Lindenbaum del sistema **S4** costituisce un'algebra di chiusura in cui l'operatore C è associato al connettivo di possibilità $\Diamond$ (nel senso che $C[\mathscr{A}] = [\Diamond\mathscr{A}]$) e quello di interno alla necessità (cosicché $I[\mathscr{A}] = [\Box\mathscr{A}]$). Il teorema di rappresentazione permetteva di concludere che le leggi di **S4** coincidono con le formule valide in ogni valutazione in algebre di chiusura e lo schema sarebbe stato esteso successivamente, soprattutto a partire da alcuni lavori di Edward Lemmon (1930-1966) del 1966, a sistemi più deboli di logica modale modificando le proprietà dell'operatore C.

Prodotto collaterale di questo risultato era una nuova dimostrazione della possibilità di interpretare la logica intuizionista nel sistema modale **S4** (come indicato da Gödel in una nota del 1933) in modo che per ogni formula $\mathscr{A}$ si ha che $\mathscr{A}$ è un teorema intuizionista se e solo se una sua traduzione $\mathscr{A}^*$ lo è in **S4**. McKinsey e Tarski oltre a dimostrare che questo era possibile per la traduzione indicata da Gödel, davano ulteriori esempi di interpretazioni di questo tipo basandosi sempre sui collegamenti tra $\Diamond$ e chiusura e sul fatto che le leggi intuizionistiche coincidono con le formule vere in ogni algebra di aperti. Era in forza di questo collegamento che risultava possibile estendere alla logica intuizionista quelle tecniche per costruire contromodelli finiti che originariamente McKinsey aveva trovato per le algebre di chiusura e aveva applicato al problema della decisione di **S4** e del più debole sistema **S2**. Sempre con metodi di questo tipo McKinsey aveva dimostrato che i connettivi $\wedge$, $\vee$, $\rightarrow$, $\neg$ sono indipendenti rispetto alla logica intuizionista (una prima dimostrazione era stata data da Wajsberg nel 1933) e aveva esteso tanto alcuni teoremi di J. Dugundji sulla non esistenza di matrici caratteristiche finite per vari sistemi modali, quanto il risultato di Gödel secondo il quale nel calcolo proposizionale intuizionista esistono infinite formule non equivalenti.

Il fatto più importante era però che il teorema di rappresentazione per le algebre di chiusura, come il suo corollario sulle algebre di Heyting, portava in primo piano ancora una volta i filtri ed in particolare – nel caso delle algebre di Heyting – i filtri *primi*. Era appunto associando ad ogni elemento dell'algebra la famiglia dei filtri primi che lo contengono che si otteneva una rappresentazione delle algebre di Heyting come algebre di aperti e d'altra parte sempre ai filtri primi si giungeva una volta che data l'algebra di Lindenbaum del calcolo intuizionista si consideravano le controparti delle teorie: come i filtri propri erano le controparti delle teorie coerenti, così erano i filtri primi ad essere le controparti delle teorie complete in senso intuizionista (per cui cioè vale la proprietà della disgiunzione). Un discorso analogo si poteva fare per le teorie nel sistema **S**4.

Furono queste tecniche che Tarski usò ancora nel 1950 quando in collaborazione con Bjarni Jónsson iniziò lo studio delle *algebre di Boole con operatori* che generalizzano tanto le algebre di chiusura quanto altre strutture che occorrono in contesto logico e matematico (per esempio la teoria della misura).

Anche qui era possibile ottenere teoremi di rappresentazione che davano monomorfismi ρ: $\mathbf{A} \rightarrow \mathbf{B}$ usando i quali era possibile analizzare gli operatori dell'algebra in termini di relazioni per ultrafiltri. Così che se, ad esempio,

$$\mathrm{C}(\rho a) = \{F \text{ ultrafiltro} \mid \exists F' \in \rho(a) \text{ tale che } FRF'\}$$

le proprietà della relazione R dipendevano dalle proprietà dell'operatore C. Si trattava di un ulteriore raffinamento dei risultati precedenti sulle algebre di chiusura che non avrebbe tardato a dare i suoi frutti nell'analisi dei calcoli proposizionali. Sarà infatti a partire da questi lavori che – per sua esplicita dichiarazione – Saul Kripke prenderà le mosse nel 1959 nella sua costruzione della semantica a mondi possibili per le logiche modali, identificando gli ultrafiltri con i mondi possibili e la relazione R con una relazione di accessibilità tra mondi. Tra i risultati più interessanti nel campo della semantica per le logiche modali di questi ultimi anni sono alcuni lavori di Robert Goldblatt, Stephen Thomason e Johan van Benthem, dedicati appunto alla *dualità* (nel senso di Stone) tra algebre con operatori e strutture di Kripke, dualità che mostra l'in-

tertraducibilità tra operazioni e relazioni tra algebre e operazioni e relazioni tra supporti.

L'applicazione di metodi algebrici non si limita però alle logiche di cui ci siamo occupati finora. A partire dai primi anni cinquanta diversi ricercatori (tra i quali in particolare Helena Rasiowa, Roman Sikorski, G. Diego, P.C. Rosenbloom) hanno studiato sistematicamente diversi sistemi proposizionali non classici: la logica minimale, i calcoli implicativi, diverse forme di logica costruttiva con negazione forte, le logiche polivalenti di Post e tante altre, ottenendo risultati che non si limitano ai teoremi di completezza e alla identificazione dei teoremi dei calcoli in esame mediante specifiche matrici caratteristiche. L'uso sistematico della teoria dei filtri in reticoli distributivi permetteva di affrontare in chiave algebrica anche risultati quali il teorema di compattezza (che nel caso classico ed intuizionista costituisce la controparte logica del fatto che lo spazio sui cui aperti (o clopen) si rappresentano gli elementi dell'algebra è compatto) i vari teoremi di interpolazione (tipo quelli di Craig e Beth di cui parleremo nel capitolo VI) come pure l'elaborazione di metodi di decisione basati sulla costruzione di contromodelli finiti.

Tra i sistemi logici più riottosi ad un'analisi algebrica adeguata risultarono paradossalmente le logiche polivalenti di Łukasiewicz e diversi sistemi di logica rilevante presentati da Ross Anderson, N. Belnap ed altri a partire dai primi anni sessanta. Un'analisi algebrica delle logiche polivalenti (come vedremo meglio nel capitolo VI) sarà iniziata nel 1959 da Chen Chung Chang mentre ancora più tarde sono le ricerche sull'algebrizzazione dei sistemi di *entailment* e della logica rilevante. La difficoltà per questi, come per altri sistemi nati dall'analisi logica della meccanica quantistica, sta nella impossibilità di definire – in mancanza di una buona nozione di condizionale – un concetto utile di algebra di Lindenbaum. Non tutti i sistemi logici ammettono un'adeguata analisi algebrica e l'estensione di questi metodi è stata in tempi recentissimi oggetto d'indagine da parte di vari autori.

Malgrado questi limiti si può dire che lo studio della logica proposizionale dal punto di vista algebrico sia oggi un fatto acquisito che ha trovato la sua canonizzazione a partire dal 1963 in un testo ormai classico di H. Rasiowa e R. Sikorski, *The Mathematics of Metamathematics* (*La matematica della metamatematica*). In quest'opera

non si trova solo una trattazione della logica proposizionale; ampio spazio è dedicato anche allo studio della logica del primo ordine sia classica che modale che intuizionista, lungo una prospettiva indicata da Mostowski e che si è rivelata estremamente illuminante nei primi anni sessanta, col sorgere della teoria dei modelli booleani per la teoria degli insiemi.

Risultati analoghi sarebbero stati ottenuti ben presto anche per la logica minimale, per quella positiva e per diverse forme di logica costruttiva con operatori di negazione forti. L'idea di fondo è quella di generalizzare il concetto di modello per la logica classica introdotto da Tarski. Esemplificando per il caso intuizionista, si tratta di considerare strutture del tipo $<D, \mathbf{H}, v>$ dove D, il dominio, è dato da un insieme non vuoto di individui, **H** è un'algebra di Heyting *completa* (in cui cioè *ogni* famiglia di elementi ha infimo e supremo) e si interpretano le costanti relazionali R del linguaggio come relazioni *in senso generalizzato* assegnando ad ogni R una funzione $v(R) : D^n \rightarrow \mathbf{H}$; ad ogni n-upla del dominio, $v(R)$ dà quindi il valore che la formula $R(x_1,..., x_n)$ prende nell'algebra una volta che ad ogni x_i sia sostituito l'i-esimo elemento della ennupla. La verità di una formula in un modello **M** così definito sarà data simulando quanto fatto nel caso proposizionale, e diremo che **M** rende *vera* la formula chiusa $\mathscr{A}$ quando il valore $\bar{v}(\mathscr{A})$ coincide con 1. Quale sia questo valore è definito al solito per induzione, facendo corrispondere ai connettivi le operazioni dell'algebra, ponendo ad esempio che

$$\bar{v}(\mathscr{A} \wedge \mathscr{B}) = \bar{v}(\mathscr{A}) \wedge \bar{v}(\mathscr{B}); \quad \bar{v}(\mathscr{A} \rightarrow \mathscr{B}) = \bar{v}(\mathscr{A}) \rightarrow \bar{v}(\mathscr{B});$$

nel caso dei quantificatori si farà ricorso a supremi e infimi ponendo che:

$$\bar{v}(\exists x \mathscr{A}) = \bigcup_{a \in D} \bar{v}(\mathscr{A}[a]) \quad \bar{v}(\forall x \mathscr{A}) = \bigcap_{a \in D} \bar{v}(\mathscr{A}[a])$$

Il passo era naturale, ma imponeva un'attenzione più sistematica agli operatori infinitari su reticoli e strutture connesse. Dopo Schröder, era stato Tarski nei suoi studi del 1935 sulle algebre di Boole complete ad aprire la strada analizzando le varie forme di distributività possibili e il loro significato. Un'attenzione analoga alle estensioni infinitarie di $\wedge$ e di $\vee$ nasceva in altri contesti, quali

ad esempio la teoria della misura in cui gli insiemi boreliani (su cui le misure sono definite) costituiscono un'algebra di Boole σ–completa, in cui cioè è possibile la riunione di famiglie *numerabili* di insiemi. Sempre nel contesto della teoria della misura, gli studi di Ulam avevano portato in primo piano le proprietà di *completezza* dei filtri (il fatto cioè che se una famiglia $\{a\}_{i \in I}$ di elementi appartiene a un filtro, anche $\bigcap_{i \in I} a_i$ vi appartiene).

Questo problema veniva a porsi anche nel caso dell'algebrizzazione della logica del primo ordine in cui ai quantificatori si facevano corrispondere operazioni infinitarie. Anche nel caso del primo ordine infatti si manteneva quel collegamento fra filtri e sistemi deduttivi (o teorie) che abbiamo visto nel caso proposizionale. Una volta costruita l'algebra di Lindenbaum del calcolo come per la logica delle proposizioni, ai quantificatori corrispondevano $\bigwedge$ e $\bigvee$ e diveniva essenziale sapere quando, data una teoria $\mathfrak{T}$ e una formula $\exists x \mathscr{A} \in \mathfrak{T}$, si disponeva di un esempio $\mathscr{A}(a)$ in $\mathfrak{T}$. È questa la proprietà di *ricchezza* delle teorie (definita per la prima volta – come avremo occasione di vedere – da Leon Henkin nel 1949) che una volta tradotta in termini di filtri dell'algebra di Lindenbaum ci porta a chiederci quando un filtro F *preserva* i supremi di una data famiglia di insiemi, vale a dire quando, dato un qualunque insieme $\{a_i\}_{i \in I}$ della famiglia, si ha che

$$\text{se } \bigcup_{i \in I} a_i \in F \text{ allora } \exists i \in I \text{ tale che } a_i \in F$$

(la preservazione degli infimi è data dualmente e afferma che

$$\text{se } \forall i \in I\ a_i \in F \text{ allora } \bigcap_{i \in I} a_i \in F).$$

Nel caso di ultrafiltri le due forme di preservazione si equivalgono e come si vede costituiscono una generalizzazione della proprietà di essere primo ed un indebolimento (relativizzato agli insiemi della famiglia) della completezza.

Il problema dell'esistenza di filtri «sufficientemente» completi con proprietà di non banalità (vale a dire né principali né degeneri) è estremamente complesso e coinvolge – come mostravano i lavori legati alle ricerche di Ulam sui cardinali misurabili – questioni profonde di teoria degli insiemi. Diverso è il caso della preservazione per famiglie vincolate in base a considerazioni di cardina-

lità. Qui Rasiowa e Sikorski ottennero un risultato (noto oggi, appunto, come *lemma di Rasiowa e Sikorski*) che afferma che, nel caso la famiglia di insiemi sia *numerabile*, è sempre possibile, per ogni $a \neq \emptyset$, trovare un ultrafiltro F che contiene a e che preserva supremi ed infimi della famiglia. Dimostrato originariamente nel 1950 utilizzando metodi topologici (in particolare il teorema di Baire) il lemma si può dimostrare più semplicemente, come indicato da Tarski, mimando la costruzione di teorie complete e ricche usata da Henkin nella sua dimostrazione del 1949 del teorema di completezza per la logica del primo ordine. Se in questo caso era l'algebra che imitava la logica, la stessa dimostrazione di Henkin, che sfruttava il teorema di Lindenbaum sulle estensioni complete, poggiava sull'analogia tra teorie e filtri e rimandava come paradigma al teorema dell'ultrafiltro di Stone. Sta di fatto che la dimostrazione algebrico-topologica del teorema di completezza per la logica classica che Rasiowa e Sikorski ottenevano dal loro lemma, costituisce l'esatto analogo di quella di Henkin anche se – non potendosi estendere il lemma a famiglie arbitrarie – pagava la maggior limpidezza concettuale con la necessità di limitarsi a teorie numerabili.

Questo parallelismo tra principi algebrici ed insiemistici e risultati metamatematici poneva in luce la possibilità di trovare equivalenti metamatematici dei principi insiemistici (quali l'assioma di scelta) e risultati in questo senso furono presto ottenuti dallo stesso Henkin, che nel 1954 provò l'equivalenza tra teorema dell'ultrafiltro e teorema di completezza per la logica del primo ordine. Nello stesso anno Dana Scott dimostrava che il teorema dell'ultrafiltro per reticoli distributivi *qualsiasi* non solo era implicato dall'assioma di scelta, ma a sua volta lo implicava. Per algebre di Boole l'esistenza di filtri massimali e filtri primi è equivalente, mentre così non è per reticoli distributivi generali; sempre Scott, nello stesso lavoro, dimostrava che l'esistenza di filtri primi per algebre di Boole implica che lo stesso vale in reticoli distributivi qualsiasi cosicché il problema diventava quello di sapere se il teorema del filtro primo implica anch'esso l'assioma di scelta.

La risposta, come vedremo, sarebbe venuta nel 1964, ma per quanto riguarda l'effettività dei risultati, ciò che emergeva chiaramente era che se innegabilmente le tecniche algebriche sono più compatte dal punto di vista concettuale, spesso esigevano principi

insiemistici non necessari in approcci alternativi di tipo sintattico. Questo valeva per diversi risultati che Rasiowa e Sikorski dimostravano, primo fra tutti quel teorema di Herbrand che era stato uno dei paradigmi della metamatematica finitista. Se da questo punto di vista le tecniche algebriche potevano sembrare a rimorchio di quelle più strettamente logiche, il lato più interessante di queste indagini non fu adeguatamente notato allora, ed era la generalizzazione stessa del concetto di modello su cui queste tecniche si basavano.

Nel caso classico, i modelli introdotti da Mostowski e dalla sua scuola erano in sostanza quelli che oggi si chiamano *modelli booleani*, strutture cioè in cui le relazioni vengono interpretate in termini di funzioni con valori in algebre di Boole complete, non necessariamente l'algebra **2** dei valori di verità. Non occorreva nessun principio insiemistico particolare per dimostrare che la logica classica è completa rispetto a questi modelli e non solo rispetto a quelli classici, in cui le relazioni hanno funzioni caratteristiche in **2**. Ciò poteva sembrare non interessante, ed infatti la nozione di modello algebrico era stata introdotta essenzialmente per trattare i modelli per logiche diverse dalla classica, per la quale si disponeva già di un'adeguata nozione di modello. Solo più tardi sarebbe emerso il senso di queste generalizzazioni, quando si sarebbero affrontate specifiche teorie (quale quella degli insiemi) e linguaggi più forti (come i linguaggi infinitari) in cui non valgono teoremi di completezza e compattezza per la nozione standard di modello. Se si vuole qualche forma di completezza occorre generalizzare il concetto di modello ed è in questa ottica che le nozioni elaborate da Rasiowa e Sikorski diventano utili.

Non è qui il caso di percorrere in dettaglio la storia di questo processo di algebrizzazione che coinvolse, a partire dalla metà degli anni cinquanta, anche lo studio dei linguaggi infinitari (prima proposizionali e poi anche con quantificatori). Ci basti ricordare come tappe fondamentali – in cui ebbero un ruolo centrale i modelli booleani generalizzati per la logica classica – i risultati di completezza ottenuti da Carol Karp nel 1964 e quelli più generali di Richard Mansfield del 1972. Queste ricerche sarebbero poi confluite nelle indagini sulla logica categoriale del primo ordine iniziate da Michael Makkai e Gonzalo Reyes nel 1976, in cui al centro dell'attenzione sono linguaggi del tipo $L_{\infty,\omega}$ (congiunzioni e

disgiunzioni arbitrarie, quantificazioni finite) ed in particolare i loro frammenti *geometrici* (contenenti solo $\wedge$, $\vee$, $\exists$) che vengono interpretati su strutture ancora più generali di quelle introdotte da Mostowski per la logica classica e intuizionista, vale a dire categorie dotate di una topologia di Grothendieck.

Non ci è possibile entrare in dettagli su questi lavori di estremo interesse che hanno ampliato il campo d'applicazione della teoria dei modelli allo studio di situazioni matematiche nuove e fondamentali. Quello che vogliamo osservare è che se gli sviluppi riguardanti la logica del primo ordine ed infinitaria di cui abbiamo parlato hanno rivelato aspetti interessanti nella teoria dei modelli, non si può nel loro caso parlare di una vera e propria *analisi algebrica* delle logiche in questione, analoga a quella condotta per i linguaggi proposizionali.

Ciò che rimaneva fuori dal raggio d'azione della teoria dei reticoli, erano le tre nozioni fondamentali che intervengono nella logica del primo ordine come in generale in quelle di ordine superiore: la quantificazione, la sostituzione di variabili e la nozione di identità vista come relazione costante con proprietà specifiche. Sulle prime, ad essere affrontato fu un tema di portata leggermente più limitata, e cioè la possibilità di dare una veste analoga a quella data alla logica proposizionale anche al *calcolo delle relazioni* (binarie) che era stato sviluppato nei lavori di Peirce e Schröder. Fu ancora Tarski che nel 1941, in un articolo intitolato *On the calculus of Relations* (*Sul calcolo delle relazioni*) tentò questa analisi introducendo il concetto di *algebra relazionale*. Un'algebra relazionale (per brevità, una **AR**) è una struttura del tipo

$$\mathbf{A} = \langle A, \wedge, \vee, -, 0, 1, \circ, \breve{.}, 1^* \rangle$$

dove $\langle A, \wedge, \vee, -, 0, 1 \rangle$ è un'algebra di Boole, $\circ$, $\breve{}$, sono operazioni binarie (il prodotto e il passaggio al converso) mentre 1^* è un elemento privilegiato di A (la *diagonale*). Le leggi che le nuove operazioni devono soddisfare non sono altro che le proprietà centrali del prodotto peirceano di relazioni, del passaggio alla relazione conversa e della relazione diagonale, la relazione che corrisponde all'identità. Come è facile immaginare, le algebre relazionali concrete di cui la teoria generale vuole dare una formulazione algebrica non sono altro che le strutture in cui A è un sottoinsieme del-

l'insieme delle parti $\mathscr{P}(I^2)$ della potenza cartesiana di un insieme I. Gli elementi di **A** sono insiemi di coppie ordinate e sono le *relazioni* di cui il dominio A è costituito. Mentre le operazioni booleane saranno le ordinarie operazioni insiemistiche, il prodotto $\circ$ di due relazioni R e S sarà dato dalla relazione

$$R \circ S = \{<u, w> \mid \exists v <u, v> \in R \text{ e } <v, w> \in S\}$$

mentre $\breve{}$ sarà l'operazione per cui

$$\breve{R} = \{<u, w> \mid <w, u> \in R\}$$

e 1* la diagonale

$$1^* = \{<u, u> \mid u \in I\}.$$

La possibilità di rappresentare *tutte* le algebre relazionali in termini concreti (così da mostrare che il concetto astratto è adeguato) fu indagata da Tarski stesso in collaborazione con Jónsson negli articoli sulle algebre con operatori che abbiamo sopra citato. Il fatto sorprendente era – come provato da Roger Lyndon nel 1950 – che non tutte le algebre di relazioni sono rappresentabili concretamente e i risultati di Tarski e Jónsson fornivano quindi rappresentazioni solo in parte concrete. Malgrado l'interesse che le algebre relazionali hanno suscitato recentemente nella *computer science*, si può dire che esse – come le *algebre proiettive* introdotte nel 1946 da C. Everett e Ulam per scopi analoghi – costituirono tutto sommato un tentativo fallito che fu poi abbandonato, anche se l'ultimo lavoro che Tarski portò a termine prima della morte in collaborazione con S. Givant, riprende l'idea di una formulazione degli assiomi per una teoria degli insiemi puramente equazionale, senza uso di variabili.

Fu ancora Tarski (in collaborazione con i suoi studenti Louise H. Chin e Federick B. Thompson) a dare inizio ad un nuovo tentativo di algebrizzazione della logica elementare nel periodo 1948-1952 collegandosi ai lavori sulle algebre relazionali e proiettive di cui si è detto. Questa volta il concetto chiave era quello di *algebra cilindrica di dimensione* α (dove α è un ordinale), una struttura

$$<A, \wedge, \vee, -, 0,1, c_\lambda, d_{\lambda\kappa}>_{\lambda,\kappa<\alpha}$$

dove al solito $<A, \wedge, \vee, -, 0,1>$ è un'algebra di Boole, ogni c_λ è

un'operazione unaria e $d_{\lambda\kappa}$ sono elementi privilegiati di A per cui valgono gli assiomi

A1 $c_\lambda 0 = 0$
A2 $x + c_\lambda x = c_\lambda x$
A3 $c_\lambda (x \cdot c_\lambda y) = c_\lambda x \cdot c_\lambda y$
A4 $c_\lambda c_\kappa x = c_\kappa c_\lambda x$
A5 $d_{\lambda\lambda} = 1$
A6 Se $\lambda \neq \kappa, \mu$ allora $d_{\lambda\mu} = c_\lambda (d_{\lambda\kappa} \cdot d_{\lambda\mu})$
A7 Se $\lambda \neq \kappa$ allora $c_\lambda (d_{\lambda\kappa} \cdot x) \cdot c_\lambda (d_{\lambda\kappa} \cdot -x) = 0$.

Non è difficile riconoscere l'origine logica delle nuove operazioni: le operazioni c_λ (dette di *cilindrificazione*) corrispondono ai quantificatori sulla λ-esima variabile, mentre i $d_{\lambda\kappa}$ (gli elementi *diagonali*) corrispondono all'equazione $(x_\kappa = x_\lambda)$; se infatti, dato un linguaggio elementare $\mathscr{L}$, si costruisce l'algebra di Lindenbaum (classica) in modo analogo a quello usato nel caso proposizionale, si avrà che

$$[\exists x_\lambda \mathscr{A}] = c_\lambda [\mathscr{A}] \qquad \text{e} \qquad d_{\kappa\lambda} = [x_\kappa = x_\lambda].$$

Se questo è il collegamento con l'aspetto linguistico della logica elementare, la dimensione semantica risulta chiara se si considera che gli esempi fondamentali di algebre cilindriche di dimensione concrete si ottengono prendendo come elementi sottoinsiemi di un'opportuna potenza I^α e definendo

$$C_\lambda X = [u \in I^\alpha \mid \exists v \in X \, u_\kappa = v_\kappa \text{ per ogni } \kappa \neq \lambda]$$
$$D_{\lambda\kappa} = \{u \in I^\alpha \mid u_\kappa = v_\lambda\}.$$

La definizione degli operatori C_λ di cilindrificazione costituisce l'aspetto nuovo delle algebre in esame e pone in luce il legame profondo fra *proiezione* lungo un asse e l'applicazione del quantificatore esistenziale, legame che Tarski aveva già applicato alla teoria descrittiva degli insiemi in un vecchio lavoro del 1930 in collaborazione con Kuratowski e che è fondamentale nello studio degli insiemi definibili.

Ciò che rendeva plausibile l'uso delle algebre cilindriche nello studio della logica del primo ordine è che le strutture per questi linguaggi hanno naturalmente associate algebre cilindriche concrete cosicché i rapporti tra aspetto sintattico e linguistico – come

nel caso proposizionale – si possono analizzare in termini di eventuali teoremi di rappresentazione. Teoremi di questo tenore (per classi particolari di algebre cilindriche, quelle *localmente finite*) furono ottenuti da Tarski e Henkin e pubblicati per la prima volta nel 1956 e risultano intertraducibili con i teoremi di completezza. Come le *algebre poliadiche* introdotte da Paul Halmos nel 1962 per lo stesso scopo, le algebre cilindriche sono state fatte oggetto di numerosi studi che ne hanno posto in luce i mutui rapporti e le proprietà algebriche, anche al di là delle applicazioni alla logica, divenendo così un capitolo interessante – anche se marginale – dell'algebra universale. Il loro uso nell'indagine logica, d'altra parte, non si può dire si sia mostrato così decisivo come quello dei reticoli nel caso proposizionale, reticoli che ancora di recente, con le *algebre diagonali* introdotte da Roberto Magari nel 1968, si sono rivelati uno strumento estremamente limpido nello studio dei fenomeni di autoriferimento collegati al teorema di Gödel.

In una prospettiva di genuina algebrizzazione, i risultati più profondi e significativi sono venuti da un altro ambito di ricerca e più precisamente dalla teoria delle categorie. Alla base di questi sviluppi sono le ricerche di William Lawvere sulla *semantica funtoriale* che fin dal loro primo apparire nel 1963 hanno portato ad una concezione nuova dell'algebra universale ed al successivo svilupparsi di una vera e propria *logica categoriale* i cui concetti base – oltre che a Lawvere – sono dovuti soprattutto ad André Joyal. In questa prospettiva, le teorie divengono particolari categorie (come nell'approccio algebrico tradizionale divenivano reticoli), le interpretazioni morfismi e gli universi di interpretazione altre categorie. Ciò si verifica non solo per la logica proposizionale (finitaria e non) ma anche per quella del primo ordine e di ordine superiore. Le categorie che possono fungere da universi sono diverse a seconda dei linguaggi che si considerano ed è la combinazione di proprietà di chiusura e di esattezza che assume il ruolo centrale. In questo contesto sono i *topos* a prendere le parti dell'universo degli insiemi – in cui dal punto di vista classico sono definite le interpretazioni – universo che a sua volta costituisce un particolare topos. In questo modo geometria, topologia generale, algebra e logica trovano un nuovo punto d'incontro in cui, diversamente che nell'approccio usuale, è la logica intuizionista a svolgere il ruolo di logica fondamentale.

W. e M. Kneale, *Storia della logica*, Einaudi, Torino 1972 [edizione originale 1962].

F. Barone, *Logica formale e logica trascendentale*, 2 voll., I. *Da Leibniz a Kant*; II. *L'algebra della logica*, Edizioni di «Filosofia», Torino 1957, 1965.

J.M. Bochenski, *A History of formal Logic*, Chelsea, New York 1970 [trad. it., 2 voll., *La logica formale*. I. *Dai presocratici a Kant*; II. *La logica matematica*, Einaudi, Torino 1972].

N.I. Stjatzkin, *Storia della logica*, a cura di Roberto Cordeschi, Editori Riuniti, Roma 1980 [edizione originale 1964].

R. Blanché, *Logica e assiomatica*, La Nuova Italia, Firenze 1968 [edizioni originali 1957-1959].

Id., *La logica e la sua storia da Aristotele a Russell*, Ubaldini, Roma 1973 [edizione originale 1970].

F. Enriques, *Per la storia della logica*, Zanichelli, Bologna 1987 [edizione originale 1922].

AA.VV., *La logica da Leibniz a Frege*, a cura di Massimo Mugnai, Loescher, Torino 1982.

AA.VV., *La logica del Novecento*, a cura di Ettore Casari, Loescher, Torino 1981.

R. Whately, *Elements of logic*, Clueb, Bologna 1987 [edizione originale 1826].

A. De Morgan, *On the Syllogism and Other Logical Writings*, Routledge e Kegan Paul, Londra 1966.

G. Boole, *L'analisi matematica della logica*, Silva, Milano 1965.

Id., *Indagine sulle leggi del pensiero su cui sono fondate le teorie matematiche della logica e della probabilità*, Einaudi, Torino 1976.

A. Del Re, *Lezioni di algebra della logica*, Tipografia della R. Accademia delle Scienze fis. e mat., Napoli 1907.

F. Gillot, *Algèbre et logique d'aprés les textes originaux de G. Boole et W.S. Jevons, avec les plans de la machine logique*, Libraire Scientifique Albert Blanchard, Parigi 1962.

P. Freguglia, *L'algebra della logica*, Editori Riuniti, Roma 1978.

T. Hailperin, *Boole's Logic and Probability*, North Holland, Amsterdam 1976.

P.E.B. Jourdain, *Selected essays on the history of set theory and logics (1906-1918)*, Clueb, Bologna 1989.

W.S. Jevons, *Lezioni di logica elementare*, a cura di G. Capone Braga, Cedam, Padova 1948 [edizione originale 1870].

Id., *The Principles of science*, Londra 1874.

J. Venn, *Symbolic Logic*, MacMillan, Londra 1881.

Id., *The Principles of empirical or inductive Logic*, MacMillan, Londra 1889.

E. Schröder, *Vorlesungen über die Algebra der Logik (Exakte Logik)*, Chelsea Publishing Co., New York 1966 [edizione originale, 1890-1910].

C.S. Peirce, *Collected Papers of Charles Sanders Peirce*, 8 voll., I-VI, Cambridge (Mass.) 1931-35, VII-VIII, id., 1958. Delle opere di Peirce è attualmente in corso un'altra edizione, della quale sono finora apparsi 5 volumi: *Writings of Charles S. Peirce. A Chronological Edition*, voll. I-V, Bloomington 1982-1990.

A.N. Whitehead, *A treatise of universal algebra*, Hafner 1960 [edizione originale 1897].

L. Couturat, *Algèbre de la logique*, Olms, Hildesheim 1965 [edizione originale 1914].

C.I. Lewis, *A survey of Simbolic Logic. The classic algebra of logic*, Dover Publications, Inc., 1960 [edizione originale 1918].

R. Wojcicki, *Theory of logical calculi*, Reidel, Dordrecht 1988.

Id., *Lectures on propositional calculi*, Ossolineum, Wroclaw 1984.

L. Henkin, J.D. Monk, A. Tarski, *Cylindric Algebras*, I, II, North Holland, Amsterdam 1971, 1985.

G. Grätzer, *Universal Algebra*, van Nostrand, Princeton 1968.

P. Halmos, *Algebraic Logic*, Chelsea, New York 1962.

J. Wolenski, *Logic and Philosophy in Lvov-Warsaw School*, Kluver Academic Publishers, Dordrecht 1983.

G. Birkhoff, *Lattice Theory*, AMS, Providence 1940.

H. Mehrtens, *Die Entstehung der Verbandstheorie*, Gestenberg, Hildesheim 1979.

L. Vercelloni, *Filosofia delle strutture*, La nuova Italia editrice, Firenze 1989.

A. Tarski, *Logic, Semantics, Metamathematics*, Hackett, Indianapolis 1983.

H. Rasiowa e R. Sikorski, *The mathematics of metamathematics*, Panstwowe Wydawnictwo Naukowe, Varsavia 1963.

CAPITOLO TERZO
LA NASCITA DELLA LOGICA MATEMATICA

1. PRELIMINARI

Come abbiamo visto, la creazione dell'algebra della logica verso la metà del XIX secolo aveva portato alla luce due fatti di estrema importanza: la possibilità di costruire una *teoria matematica* della logica e l'esistenza di un legame profondo, anche se difficile da esplicare, fra logica, matematica e filosofia. Mentre però il primo fatto, la costruzione di una teoria algebrica della logica, non tardò ad essere universalmente accettato (anche se, in verità, non sempre in modo entusiastico), il secondo, l'esistenza di strette connessioni fra logica e matematica, divenne sì un tema di cui si riconobbe ben presto l'importanza, ma del quale solo lentamente emersero le conseguenze meno immediate e più problematiche. La storia della logica nella seconda metà dell'Ottocento è in larga misura la storia di come andò lentamente prendendo forma, e quindi configurandosi stabilmente, questa nuova concezione dei rapporti fra logica e matematica.

Sulle prime il fatto più rilevante e più sottolineato rimase pur sempre, come era stato per gli stessi creatori dell'algebra della logica, l'universalità d'*applicazione* del metodo simbolico e, quindi, del procedere matematico. Era soprattutto in questa prospettiva di matematica applicata che si valutava l'importanza dell'algebra della logica: si trattava di un campo d'applicazione di tipo nuovo, diverso dai tradizionali, che poneva in rilievo come il dominio d'impiego della matematica non si limitasse al mondo delle quantità, ma si estendesse ad ambiti non legati al mondo fisico, quantitativo, quale il mondo dei pensieri. L'algebra della logica era così una teoria matematica del pensiero: ma più che di prossimità, di legami fra il concetto di numero, di quantità e grandezza, con con-

cetti puramente logici, si trattava di un'applicazione giustificata dalla universalità del solo procedere simbolico, che in questo modo veniva ad assumere un vero e proprio ruolo fondante.

Veniva così prendendo forma, soprattutto in Germania – come si è visto – una nuova concezione della matematica e della sua struttura logica, una concezione con decisi caratteri formalistici (in senso diverso, beninteso, da quello che sarà il formalismo hilbertiano) che vedeva soprattutto nella capacità della mente umana di astrarre e fare un uso appropriato dei simboli il fondamento dell'intera attività matematica e in questa prospettiva la logica e la sua algebrizzazione non avevano alcun ruolo fondante ma costituivano solo un esempio, di particolare significato sistematico, di questo procedere simbolico. L'unico vero problema riguardava l'adeguatezza delle astrazioni e delle simbolizzazioni su cui si innestava il procedere matematico: che garanzie c'erano che alla regolarità data dalle leggi matematiche e dalle connessioni simboliche corrispondessero regolarità fra gli oggetti e le relazioni concrete designate? In altri termini, si trattava di indagare che cosa giustificasse la riconosciuta adeguatezza e l'innegabile successo del procedere simbolico, quali fossero le *condizioni* che rendevano possibile la sua applicazione. A questo punto il problema non riguardava più né la logica né la matematica ma la costituzione interna della mente umana, e la questione si spostava decisamente su un terreno filosofico tradizionale: *era questo il problema dei fondamenti.*

Ora, le soluzioni filosofiche che si dettero al problema furono varie: si andava da forme di kantismo più o meno ortodosse a posizioni realistiche o decisamente empiristiche. Carattere comune di tutte queste soluzioni rimaneva però pur sempre il tentativo di agganciare la concreta struttura delle varie teorie matematiche a fatti del mondo esterno mediante il ricorso a capacità della mente umana che giustificassero la possibilità di rendere a livello simbolico queste molteplici situazioni esterne. Queste capacità innate (o acquisite, a seconda del contesto filosofico generale) dovevano giustificare la correttezza delle teorie matematiche, la loro applicabilità, il loro significato. Sia che, empiristicamente come faceva Mill, si vedesse nei simboli numerici (o geometrici) un nome per fatti del mondo esterno, sia che, come Schröder o Helmholtz, i simboli denotassero astrazioni da fatti fisici o psicologici, rimaneva tuttavia il problema di giustificare la possibilità stessa di queste

astrazioni o denotazioni e, soprattutto, di trovare le condizioni che spiegassero la concordanza fra quanto aritmetica, Analisi, geometria, affermavano a livello simbolico con quanto si sosteneva essere il substrato concreto, fisico o psicologico di queste simbolizzazioni. Non può stupire quindi che in questa prospettiva i legami fra logica e matematica non fossero che analogici, giustificati da comuni capacità della mente umana che si ponevano come condizioni della esistenza stessa tanto di logica che di matematica.

Questa concezione, con tutte le svariate articolazioni sulle quali non è qui necessario insistere, se da una parte rifletteva l'atmosfera in cui si era sviluppata l'algebra della logica, sembrava dall'altra in grado di rendere conto dei grandi mutamenti che la pratica matematica andava registrando in quel torno di secolo. La scoperta delle geometrie non euclidee e il loro diffondersi a cominciare dai grandi lavori di Riemann, Helmholtz, Lie, Klein e Beltrami attorno al '68, l'accentuato processo di riorganizzazione dell'analisi con le contemporanee definizioni di numero reale di Weierstrass, Cantor e Dedekind nel '72, avevano sostanzialmente segnato la fine del più che millenario dominio della geometria sull'intero *corpus* della matematica: sotto gli urti convergenti delle geometrie non euclidee da un lato e delle definizioni puramente aritmetiche di numero reale dall'altro, l'univocità e necessità dei concetti geometrici diventavano di fatto insostenibili. Se per la pratica matematica questo significava un allargamento degli orizzonti, dal punto di vista dei fondamenti tutto ciò non poteva non far sorgere serie perplessità.

Sino ad allora, il punto di riferimento per la quasi totalità dei concetti matematici era stata la geometria; era dall'«osservazione» delle grandezze geometriche che erano nati i concetti base dell'Analisi, quale la nozione di continuo, ed era la geometria il settore del sapere matematico che da sempre aveva avuto la struttura logica più chiara e articolata, offrendo simultaneamente un *contenuto descrittivo* chiaro ed una *organizzazione logica* agli enunciati matematici. Negare o «perdere» questo riferimento, significava abbandonare la fonte più sicura di conferme intuitive e porsi nella necessità di cercare nuovi fondamenti per l'intero edificio matematico. La costruzione di un concetto puramente aritmetico del continuo per opera di Weierstrass, Cantor e Dedekind aveva dato l'avvio a un totale cambiamento di rotta che portava in primo pia-

no, al posto della geometria, l'aritmetica e il sistema dei numeri naturali come base per analizzare concettualmente *anche* l'intuizione geometrica. Questo mutamento aveva un profondo significato in quanto segnava un netto svincolarsi delle interpretazioni dei fatti matematici dal concetto intuitivo di spazio, orientando verso nozioni almeno apparentemente più semplici e dirette, quale quella di numero naturale: nozione quest'ultima che sembrava decisamente immune da ogni difficoltà logica, legata, come sempre era stata ritenuta, alla comunissima operazione del contare.

In tal modo sembrava possibile recuperare l'antica certezza: se infatti, via l'aritmetizzazione, l'Analisi poteva essere fondata sul concetto di numero naturale, modulo la «semplice» assunzione della possibilità di astrarre e considerare come entità numeriche sequenze fondamentali (con Cantor) o sezioni (con Dedekind), la stessa costruzione di geometrie non euclidee si poteva ricondurre a fondamenti aritmetici via una opportuna traduzione analitica. I lavori di Riemann e Helmholtz, nel secondo con riferimenti fisici più espliciti che nel primo, mostravano infatti come la nozione intuitiva di spazio si potesse sussumere sotto il concetto analitico di varietà *n*-dimensionale ricevendo quindi una fondazione aritmetica. Naturalmente ciò non permetteva una discriminazione completa tra i diversi tipi di geometrie, ma era proprio questo carattere comune e unitario che costituiva una garanzia. Per l'ulteriore discriminazione occorrevano ipotesi (Riemann) o fatti (Helmholtz) aggiuntivi che fossero suggeriti non più dalla semplice intuizione, ma piuttosto dall'esperienza o da particolari necessità teoriche; ma in questo modo, si riteneva, veniva enucleato il substrato comune alle varie geometrie, substrato che, pur necessitando nei vari casi di aggiunte specifiche, poteva pur sempre costituire una fondazione comune.

Per la prima volta nella storia, dunque, era il continuo definito aritmeticamente che fondava quello geometrico e nell'*Ausdehnungslehre* (*Teoria dell'estensione*) di Herman Grassmann (1809-1877), la nozione intuitiva di spazio risultava caso particolare di un concetto di carattere astratto, quello di spazio *n*–dimensionale, definito in termini algebrici. Se tramite l'assunzione della continuità come caratteristica fondamentale della nozione di spazio era possibile ricondurre geometria ad Analisi e questa a sua volta si poteva ricondurre a concetti aritmetici, la base di tutta la matematica era

allora data dai numeri naturali. Ebbene, su che cosa si fondavano questi, a loro volta? Questa domanda, anche se naturale da un punto di vista astratto, storicamente non si pose in modo altrettanto ovvio. La concezione «formalista» cui abbiamo sopra accennato sembrava rendere superflua e irrilevante una simile questione. Il numero naturale non era che la trascrizione simbolica della semplice attività del contare e del numerare, fatto questo così primitivo e semplice da non richiedere ulteriore analisi; se specificazioni erano necessarie, si trattava di giustificazioni di carattere psicologico riguardanti le condizioni che rendevano possibile questa operazione. È su questa linea che si situano le fondazioni dell'aritmetica di Helmholtz, Schröder, Hankel, Thomae e moltissimi altri (a parte ovviamente eventuali differenziazioni *specifiche* fra i vari autori): se Helmholtz ricorreva alla successione temporale per spiegare l'ordinamento dei naturali, ponendosi così in una linea di pensiero tipicamente kantiana (seppure allargata), Schröder ricorreva alle strutture psicologiche generali che giustificano i procedimenti astrattivi.

Ma allora questo riproponeva il tema centrale già enunciato: punto di partenza comune rimaneva sempre il ricorso a teorie filosofiche generali che fornissero le *condizioni* per l'esistenza di particolari procedimenti concettuali. E l'aritmetica a sua volta diveniva un'attività simbolica resa possibile (ed applicabile al mondo esterno) da particolari proprietà della mente umana che ne garantivano l'aggancio con il mondo intuitivo. Tuttavia, anche se queste posizioni potevano sembrare plausibili da un punto di vista estremamente generale, la loro debolezza diveniva manifesta una volta che, dalla generalità, si passasse alla considerazione di fatti più specifici.

Come non tardarono a realizzare Frege e Dedekind, la lacuna fondamentale di queste teorie stava nella loro assoluta incapacità di giustificare quello che era il principio fondamentale alla base di ogni concreto ragionamento aritmetico, vale a dire il principio di induzione. Era questo principio che giustificava le definizioni stesse delle operazioni aritmetiche di somma e prodotto e che permetteva la formulazione e dimostrazione di teoremi aritmetici generali. Qualunque oggetto intuitivo si correlasse ai simboli numerici, qualunque assunzione generale si facesse sui processi psicologici soggiacenti alla costruzione dell'aritmetica, tutto ciò non poteva

giustificare il ragionamento per induzione in modo effettivamente definito e soddisfacente e porre in luce il carattere distintivo dell'operazione del contare, ossia il suo procedere per iterazione. Questa obiezione colpiva nella sua essenza tutto un modo di concepire la fondazione della matematica. Se l'aritmetica doveva essere la solida base del sapere matematico, dell'Analisi in particolare, non poteva certo essere sufficiente la certezza «morale» data dalla convinzione che dietro al procedere aritmetico esistessero processi mentali che lo giustificavano. Non era possibile spiegare in questo modo la validità del principio di induzione: il ricorso alla struttura temporale o ad altri fatti empirici (fossero essi psicologici o fisici) non poteva dare alcuna certezza ma, al più, una mera plausibilità empirica e non era in grado d'altra parte di mostrare se e in quali altri contesti procedure induttive fossero utilizzabili, quali dovessero essere le caratteristiche dei domini cui fossero applicabili numerazioni e definizioni per induzione.

La soluzione che Frege e Dedekind proposero si volgeva in tutt'altra direzione: ben lungi dal cercare fatti psicologici, leggi della mente o poteri taumaturgici dell'attività simbolica, essi, con modalità differenti, ma simili nelle motivazioni fondamentali, si volsero piuttosto all'enucleazione delle caratteristiche *logiche* che i numeri naturali dovevano possedere per giustificare a *livello deduttivo* le proprietà che ad essi l'aritmetica assegnava; sicché l'analisi condotta da Frege e Dedekind portava in primo piano l'intelaiatura puramente logico-deduttiva dell'aritmetica. In questo modo i numeri naturali non erano più *simboli* introdotti per rendere conto di fatti psicologici o fisici più o meno chiaramente intesi, bensì concetti, *costrutti logici*, definiti in modo tale da permettere la derivazione, per via puramente deduttiva, delle loro proprietà e il principio di induzione, ben lungi dall'essere un principio inspiegabile se non col ricorso a fatti extramatematici, diveniva una delle proprietà definitorie dello stesso concetto di numero. Veniva realizzandosi così un nuovo tipo di rapporto fra logica e matematica: la logica diveniva la base stessa della matematica. Il concetto di numero naturale risultava completamente determinabile in base a concetti puramente logici o, nel caso di Dedekind, insiemistici generali; era cioè un costrutto ottenuto ricorrendo a poche nozioni logiche fondamentali. È principalmente in questo lavoro e nella continua insistenza sulla ne-

cessità di esplicitare le caratteristiche logiche dei concetti matematici da definirsi, piuttosto che ricercare le condizioni che rendono possibili ipotetici atti mentali, è proprio qui, dicevamo, che sta il carattere rivoluzionario dell'opera di Frege e di Dedekind. Dopo di essi la fondazione della matematica non poteva più configurarsi come ricerca delle condizioni psicologiche in grado di giustificare determinati procedimenti intellettuali: sono i concetti stessi e le loro mutue relazioni che vengono in primo piano, non più strutture dell'esperienza, ma strutture, vincoli logici, tra concetti.

È in questa mutata atmosfera che si situa in effetti anche la costruzione cantoriana della *teoria* degli insiemi e prende forma la «revisione» moderna che dell'assiomatica fa Hilbert. Si deve anche aggiungere che mai come in questo periodo si assiste alla reciproca incomprensione fra autori che sostanzialmente, a parte le specifiche differenze, lavorano in definitiva a uno stesso, grandioso progetto. Possiamo assumere ancora Frege come esemplare in questo senso: egli polemizza con Dedekind perché non vede la (allora) sostanziale equivalenza di teoria degli insiemi e logica (che, si noti bene, era Dedekind a *esplicitare* in modo consapevole e non, paradossalmente, Cantor); rimprovera, senza dubbio correttamente, a Cantor precise deficienze logiche, e ciò senza tener conto da una parte che egli stesso aveva riconosciuto che una rigorosa sistemazione logica segue e non accompagna la «creazione» matematica, dall'altra precludendosi una più approfondita disamina dei rapporti fra i due modi, il suo e quello di Cantor, di affrontare le questioni di fondazione; ha infine polemiche accese con Hilbert, perché gli sfugge completamente che la nuova concezione hilbertiana dell'assiomatica è in fin dei conti una conferma della situazione nuova che egli stesso aveva contribuito a determinare. D'altra parte anche Dedekind, Cantor e Hilbert trascurano e non comprendono le esigenze, oltre che le innovazioni di Frege (salvo forse Dedekind); sfugge loro il fatto che nel tentativo di Frege si potesse trovare un riferimento comune alle varie tendenze specifiche, o comunque un utile strumento di discussione e apertura, non di «lotta». E a conferma che questa fosse effettivamente la situazione, che qualcuno ha definito «dolorosa», appare non casuale che la scoperta delle antinomie coinvolga in un unico «crollo» i sistemi di Dedekind, di Frege, di Cantor (da questo punto di vista

Hilbert va considerato a parte). Ma riprendiamo brevemente il nostro discorso.

Nelle mani di Frege la stessa algebra della logica veniva dunque ad assumere un carattere nuovo. La *Begriffsschrift* fregeana, accanto ai potenziamenti tecnici che vedremo, possiede una caratteristica di eccezionale importanza: è una logica che può applicarsi *direttamente* e in modo uniforme agli enunciati matematici, un algoritmo duttile (malgrado le enormi pesantezze notazionali) che segue passo passo, senza lacune, lo snodarsi delle dimostrazioni; con Frege l'aritmetica cessa di essere una scienza caratterizzata in modo vago per divenire una vera e propria *teoria* con una netta struttura logica. Discorso analogo si può fare per Dedekind, che con Frege divide il merito d'aver individuato un «sistema d'assiomi» per la teoria dei numeri naturali; sistemi d'assiomi che riceveranno con Peano una veste «linguistica» più diretta, ma che comunque – e qui sta un vero punto di diversa impostazione – si diversificano essenzialmente per le diverse concezioni «assiomatiche» di questi autori, specie se confrontate con la versione moderna, hilbertiana, dell'assiomatica.

Non abbiamo qui voluto dare una descrizione compiuta e conseguente dei grossi «fatti» che cambiano il volto della filosofia della matematica nel secondo Ottocento; abbiamo semplicemente tentato di fornire al lettore un filo conduttore generale per le pagine che seguono, nelle quali è stato ovviamente necessario operare separazioni «artificiose» fra autore e autore, anche se ridotte al minimo, sicché potrà anche avvenire, ad esempio, di trovare il pensiero di uno stesso autore analizzato sotto paragrafi diversi. Quello che più ci preme sottolineare in questa sede introduttiva è comunque la nuova prospettiva che andava emergendo dai lavori degli autori sopra citati. Dopo un lungo periodo di incertezza, prendeva forma tutta una nuova concezione dei rapporti fra logica e matematica. Ben lungi dal presentarsi come ramo dell'algebra, esempio particolare di applicazione del procedere simbolico, la logica si rivelava uno strumento incomparabile per dipanare i significati stessi delle principali nozioni matematiche; il rapporto diveniva più organico e stretto non più a livello di *metodi*, ma a livello di *concetti*. Si otteneva così una sorta di sviluppo continuo che dai concetti logici generali, mano a mano determinandosi, generava concetti aritmetici, analitici, ecc.

La logica si presentava quindi, anche se con l'intermediario – allora non da tutti compreso – della teoria degli insiemi, come il tessuto connettivo dell'intero edificio matematico. In questa prospettiva, il problema dei fondamenti da questione vagamente empirico-psicologica si era trasformato nel problema ben più rilevante della *individuazione della struttura logica dei concetti base, dei mutui rapporti fra essi e dei loro legami di definibilità*. Teoria degli insiemi e logica (ripetiamo che i due campi erano allora, in gran parte almeno, coincidenti) si presentavano così come le basi più «vere» e adeguate della nuova pratica matematica, una pratica non più legata in modo esclusivo al mondo esterno e alle intuizioni che con esso la pongono in contatto, ma libera attività creatrice del pensiero umano.

2. L'ARITMETIZZAZIONE DELL'ANALISI E I SUOI SVILUPPI

Si è già ripetutamente indicata come elemento distintivo nell'impostazione della ricerca matematica del primo Ottocento l'accentuata esigenza di rigore nella sistemazione dell'Analisi, esigenza che si era espressa con le precisazioni che autori quali Gauss, Abel, Cauchy, Bolzano, ecc. avevano dato di taluni concetti fondamentali, primo fra tutti il concetto di limite. A questa esigenza di basi concettuali più chiare i matematici erano stati condotti dallo sviluppo stesso dell'Analisi, che allargando il dominio di applicazione della teoria aveva imposto un'indagine più raffinata degli ambiti di validità dei risultati che si cercava di generalizzare portando a distinzioni, fusioni e modificazioni di concetti classici (un esempio per tutti: si pensi alla distinzione fra continuità e continuità uniforme). Tuttavia, proprio approfondendo l'esame di tali precisazioni, analizzandole minutamente nei loro presupposti più o meno esplicitamente enunciati, risultò ben presto chiaro agli analisti della seconda metà dell'Ottocento che il fondamentale concetto di *numero reale* (presupposto ad esempio in modo stringente nel celebre criterio di convergenza di Cauchy) non aveva ancora ricevuto una definizione rigorosamente «aritmetica»; esso veniva in generale assunto come fondato su intuizioni di tipo geometrico (in particolare, come rapporto fra grandezze geometriche) strettamente connesse, in modo sostanzialmente circolare, con

l'altrettanto vaga nozione di continuità (e in particolare di continuità geometrica) di per sé considerata «evidente».

In senso stretto, per aritmetizzazione dell'Analisi si può allora intendere l'affrancamento da questa «servitù» geometrica nella stessa definizione di numero reale e quindi, in una più ampia accezione, la conseguente edificazione dell'Analisi su basi che non rimandano più a intuizioni geometriche inanalizzate, ma sono esplicitate in termini di oggetti e processi *aritmetici* elementari. Agli analisti della seconda metà dell'Ottocento spetta appunto il merito di aver decisamente affrontato il compito di stabilire una rigorosa teoria dei numeri reali, con un atteggiamento che si pone in «naturale» e diretta continuazione di quell'esigenza di rigore con la quale si era aperto il XIX secolo matematico, e con esiti che andarono ben al di là del puro ambito analitico dal quale questa impresa era stata sollecitata e nel quale troverà una soddisfacente soluzione.

A parte certe precisazioni che a rigore andrebbero fatte ma che in questa sede possiamo tralasciare, intendiamo l'insieme R dei numeri reali come costituito dall'insieme Q dei numeri razionali (anche qui, a parte certe precisazioni: le frazioni) *e* dall'insieme I dei numeri irrazionali. Dato che – come vedremo – punto di partenza per le ricerche sui reali era l'insieme Q (gli elementi dei quali si sapevano già ricondurre con metodo standard ai numeri naturali, si sapevano cioè definire in termini di naturali) è chiaro che il problema della fondazione rigorosa dei numeri reali si convertiva in quello di dare una definizione soddisfacente e non obiettabile del concetto di numero irrazionale; sicché le teorie che più avanti brevemente esporremo venivano appunto qualificate come «teorie degli irrazionali».

Un numero reale veniva concepito in generale come rapporto fra grandezze, in particolare fra grandezze geometriche, fra segmenti. Era quindi la sola geometria che per così dire poteva garantire l'esistenza di un dato numero, in diretta dipendenza dalla possibilità o meno di effettuare, sulla base degli assiomi della geometria stessa, determinate costruzioni di segmenti che rispetto a una data unità possedessero come misura il numero in questione. Ora è noto, ad esempio, che non è possibile costruire (con riga e compasso), nella geometria euclidea, un segmento di lunghezza poniamo $\sqrt[3]{2}$ o π; eppure si è convinti, e ciò proprio per ragioni di

continuità, dell'*esistenza* di tali segmenti, e quindi di tali misure, di tali numeri. Nel caso di $\sqrt[3]{2}$ il segmento in questione risolverebbe il problema di Delo della duplicazione del cubo: orbene, se lo spigolo del cubo aumenta con continuità da 1 a 2, il volume del cubo aumenterà con continuità da 1 a 8 e a un certo momento dovrà quindi assumere proprio il valore 2.

Ciò mostra da una parte l'immediata connessione del problema di una caratterizzazione dell'insieme R col classico problema della continuità (geometrica); e d'altra parte rende evidente come sia illusoria e circolare la pretesa di definire in generale un numero reale come rapporto di segmenti. La situazione è presentata con particolare evidenza e incisività da Dedekind[1] che col saggio *Stetigkeit und irrationale Zahlen* (*Continuità e numeri irrazionali*) del 1872 offrirà come vedremo una originale soluzione al problema. Fin da quando, nel 1858, egli era docente al politecnico di Zurigo «...discutendo la nozione del tendere di una grandezza variabile a un valore limite, e in particolare nella dimostrazione del teorema che ogni grandezza variabile che aumenta con continuità, restando però limitata, deve certamente tendere a un limite dovevo ricorrere», afferma Dedekind nella citata opera, «all'intuizione geometrica». Se anche questa procedura risulta essere didatticamente efficace, certamente «questa forma d'introduzione del calcolo differenziale non può pretendere di essere scientifica...». Il calcolo differenziale veniva appunto presentato come quella branca della matematica che tratta di grandezze continue (e di variazioni continue di grandezze) senza che tuttavia si fosse mai ritenuto possibile o necessario dare una definizione o una soddisfacente spiegazione della continuità. «Anche le esposizioni più rigorose del calcolo differenziale,» prosegue Dedekind, «non basano le loro dimostrazioni sulla continuità bensì... o fanno ricorso a nozioni geometriche o comunque suggerite dalla geometria, oppure dipendono da teoremi

[1] Julius Wilhelm Richard Dedekind nacque a Brunschweig il 6 ottobre 1831 e ivi morì il 12 febbraio 1916. Compì i suoi studi a Gottinga e nel 1854 divenne docente nella locale università. Nel 1858 insegna Analisi al politecnico di Zurigo e nel 1862 si trasferisce come ordinario al politecnico della sua città natale, ove compie tutta la restante carriera accademica. Autore di molte importanti memorie sulla teoria dei numeri (Dedekind è il creatore della teoria degli ideali) sulla geometria algebrica, ecc., è noto tuttavia soprattutto, in ambito non matematico, per i due saggi *Stetigkeit und irrationale Zahlen* del 1872 e *Was sind und was sollen die Zahlen?* del 1888, le sole due opere che di fatto ci interesseranno nelle pagine seguenti.

che non sono mai stati stabiliti in maniera puramente aritmetica.» Da queste considerazioni scaturisce il compito che si presenta inderogabile per l'analista rigoroso (e che Dedekind assume come scopo del suo saggio): quello di giungere a scoprire la «vera origine» della continuità «negli elementi dell'aritmetica, assicurando nel contempo una definizione reale dell'essenza della continuità», sicché essa, spogliata da ogni immagine intuitiva e da ogni inutile riferimento geometrico, possa costituire il punto di partenza per effettive «deduzioni corrette». Il problema è quindi quello di analizzare l'*intuizione* del continuo e sostituirvi una definizione concettualmente determinata che non costringa, nel corso della dimostrazione, a passaggi non controllabili sul piano logico e tale quindi che non risulti determinato l'ambito di validità. Così come l'analisi dei fatti centrali riguardanti lo spazio era stata possibile ricorrendo al concetto analitico di varietà, anche la continuità andava ridotta in termini aritmetici.

Da questo punto di vista dunque tutta la revisione critica legata al problema della definizione dell'irrazionale può essere riguardata come il tentativo, che fu come vedremo effettivamente coronato da successo, *di sostituire al continuo geometrico il continuo «aritmetico»*. Su questo punto torneremo comunque più avanti (par. 2.4), dopo aver brevemente esposto le principali «teorie degli irrazionali» proposte dagli analisti di questo periodo.

Decisivo impulso a quest'opera di ricostruzione diede Karl Weierstrass.[2] Egli era venuto esponendo nelle sue lezioni all'università di Berlino una propria teoria degli irrazionali che fu successivamente presentata dal suo allievo Ernst August Martin Kossak (1839-92), e alla quale si ispirò per la propria versione un altro suo più celebre allievo, quel Georg Cantor di cui dovremo

[2] Karl Weierstrass nacque a Ostenfelde il 31 ottobre 1815. Fu avviato dal padre alla carriera giuridica e studiò giurisprudenza e scienza delle finanze a Bonn fino al 1834. Nel frattempo però egli si era applicato privatamente allo studio della matematica, e dopo aver vinto nel 1841 una cattedra liceale di questa materia, si dedicò completamente ad essa, insegnando in varie scuole. Pubblicò nel 1854 e nel 1856 fondamentali memorie sul «Journal de Crelle» che gli valsero l'assegnazione di una laurea *ad honorem* da parte dell'università di Königsberg. Nel 1864 venne chiamato a Berlino come professore ordinario. Divenne qui un caposcuola soprattutto per quanto riguarda la teoria delle funzioni ellittiche e abeliane, nonché, in generale, per l'intransigente ricerca di rigore nella sistemazione dell'Analisi. Morì a Berlino il 19 febbraio 1897 (dopo che l'accademia di Berlino aveva iniziato già dal 1894 la pubblicazione della sua *opera omnia*).

occuparci lungamente in un prossimo paragrafo. L'importanza del problema era nel frattempo divenuta centrale per lo stesso progresso dell'Analisi, sicché, con una contemporaneità non rara nella storia della scienza, nel 1872 apparvero almeno quattro lavori che proponevano sostanzialmente tre tipi diversi di soluzione. Si tratta del volume *Die Elemente der Arithmetik* (*Gli elementi dell'aritmetica*), dove Kossak esponeva appunto la teoria del maestro, che veniva anche presentata nello stesso anno dal francese Charles Meray (1835-1915) nel suo *Nouveau précis d'analyse infinitésimale* (*Nuovo compendio d'analisi infinitesimale*). Una propria soluzione ispirata alla teoria di Weierstrass, ma secondo le sue stesse parole «solo esteriormente simile» a quella, presenta succintamente Cantor in *Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen* (*Sull'estensione di una proposizione dalla teoria delle serie trigonometriche*); Cantor darà un'ampia esposizione della propria teoria undici anni più tardi nell'articolo *Über unendliche lineare Punktmanningfaltigkeiten* (*Sugli insiemi infiniti lineari di punti*). La terza teoria è infine quella presentata da Dedekind ed è contenuta nel saggio sopra citato *Continuità e numeri irrazionali*. Va ancora ricordato che sempre nel 1872 Eduard Heine (1821-82), nell'articolo *Die Elemente der Funktionenlehre* (*Gli elementi della teoria delle funzioni*) prospetta una soluzione sostanzialmente coincidente con quella di Cantor ma ottenuta indipendentemente da questi (sicché si parla anche di teoria di Heine-Cantor). Tale teoria fu ripresa nel 1877 da Rudolf Lipschitz (1832-1903) nelle *Grundlagen der Analysis* (*Fondamenti dell'Analisi*). Del 1886 è infine l'*Introduction à l'étude des fonctions d'une variable* (*Introduzione allo studio delle funzioni di una variabile*) ove Jules Tannery (1848-1910) presenta una teoria di tipo dedekindiano e che Dedekind stesso riconobbe essere stata ottenuta in modo indipendente dallo studioso francese (sicché si parla anche di teoria di Dedekind-Tannery). Si può dire che dalla fine del secolo è divenuto usuale far precedere ogni trattato d'Analisi da una teoria dei numeri reali esposta in generale o nella forma dedekindiana o, più usualmente, nella forma cantoriana. Illustriamo ora brevemente le tre definizioni.

2.1 *Definizione di Weierstrass*

Questi parte dalla considerazione di insiemi infiniti $\{a_\nu\}$ di razionali positivi che soddisfino la condizione che tutte le somme di un numero *finito* di loro elementi siano complessivamente limitate, si mantengano cioè costantemente minori di un dato intero positivo n, arbitrario ma prefissato. A ognuno di tali insiemi viene «associato» un numero b e si dimostra quindi che per tali numeri possono definirsi le usuali relazioni di «uguale», «maggiore» e «minore», in termini di proprietà degli insiemi cui essi sono associati: con lo stesso procedimento è possibile definire tra tali numeri le usuali operazioni di somma, prodotto ecc. Il momento generatore nel metodo di Weierstrass è evidentemente quello della *formazione di somme*; va esplicitamente osservato che tale momento riguarda sempre e soltanto un numero *finito* di elementi dell'insieme considerato e che il numero b associato all'insieme $\{a_\nu\}$ (o, come anche diremo, *determinato* da quell'insieme) *non* viene definito ponendolo uguale alla sommatoria sull'insieme stesso ($\Sigma_\nu a_\nu$), e ciò proprio per evitare quella circolarità di cui sopra si parlava, per evitare cioè di definire un numero tramite se stesso. Cantor attribuisce esplicitamente a Weierstrass il merito di essere stato il primo a comprendere la centralità di questo passaggio e di aver così evitato una definizione circolare.

2.2 *Definizione di Cantor*

Oltre ad avere una similitudine solo esteriore con la teoria di Weierstrass, Cantor ritiene che la propria teoria abbia su quella il vantaggio di «adattarsi in modo più immediato» al calcolo; tale vantaggio egli rivendicherà alla propria sistemazione anche nei riguardi della teoria di Dedekind. Più che da insiemi non strutturati Cantor parte da successioni infinite di numeri razionali positivi, che caratterizza con una condizione diversa da quella di Weierstrass. Precisamente egli richiede che data una tale successione $\{a_v\}$ si abbia che la differenza $a_{\nu+m} - a_\nu$ fra due suoi elementi qualsiasi diminuisca indefinitamente al crescere di ν, qualunque sia l'intero positivo m, o altrimenti detto che, co-

munque prefissato un ε positivo arbitrario «piccolo a piacere», risulti in valore assoluto

$$| a_{\nu+m} - a_\nu | < \varepsilon$$

per $\nu \geq n_1$ e m intero positivo qualunque. Con le precisazioni che seguiranno, si può esprimere la cosa con la scrittura più concisa

$$\lim_{\nu \to \infty} | a_{\nu+m} - a_\nu | = 0, \text{ per } m \text{ arbitrario.}$$

Altrimenti detto, solo per un numero *finito* di termini di una successione $\{a_\nu\}$ siffatta tale differenza risulta maggiore o uguale a ε. Successioni che soddisfano questa condizione vengono dette da Cantor *successioni fondamentali* (in terminologia moderna: successioni convergenti, o successioni di Cauchy); a ognuna di esse viene «associato» un numero b che può essere razionale (si costruiscono facilmente successioni fondamentali che hanno «limite» razionale) o irrazionale (e si introducono così questi «nuovi numeri»). Sulla base delle usuali relazioni di «uguale», «maggiore» e «minore» opportunamente estese alle successioni fondamentali $\{a_\nu\}$ e $\{a'_\nu\}$, si definiscono analoghe relazioni fra i numeri b e b' ad esse associati. Con lo stesso procedimento, alla somma $\{a_\nu + a'_\nu\}$ e al prodotto $\{a_\nu \cdot a'_\nu\}$ di due successioni fondamentali $\{a_\nu\}$ e $\{a'_\nu\}$ si associano rispettivamente i numeri $b + b'$ e $b \cdot b'$ che rappresentano la somma e rispettivamente il prodotto dei due numeri determinati da quelle successioni.

A questo punto Cantor ritiene di poter *dimostrare* rigorosamente che se b è il numero associato alla successione fondamentale $\{a_\nu\}$ allora si ha

$$\lim_{\nu \to \infty} \{a_\nu\} = b.$$

Si deve porre attenzione al fatto che, come per Weierstrass, il numero b *non* viene definito come limite della successione fondamentale $\{a_\nu\}$, altrimenti si incorrerebbe ovviamente nella circolarità più volte denunziata. Le cose stanno piuttosto, dice Cantor, esattamente al contrario: «... attraverso la nostra definizione il concetto b è pensato e costruito con tali proprietà e relazioni con i

numeri razionali che da esso può essere dedotto con evidenza logica la conclusione:

$$\lim_{\nu \to \infty} \{a_\nu\} \text{ esiste ed è uguale a } b».$$

Resta da mettere in evidenza un aspetto della teoria di Cantor che potrà essere collocato nella sua giusta luce nel seguito di questo capitolo (si veda il paragrafo 3). Una volta avviato col metodo cantoriano sopra descritto il processo di introduzione di «nuovi numeri» (gli irrazionali) per mezzo delle successioni fondamentali di numeri razionali, è chiaro che sarà possibile iterarlo considerando successioni fondamentali i cui elementi siano numeri determinati da successioni fondamentali di numeri razionali: se conveniamo di chiamare questi ultimi «del primo ordine», chiameremo quelli «del secondo ordine». Iterando successivamente il procedimento avremo successioni fondamentali di elementi del secondo ordine cui saranno associati numeri del terzo ordine, e così via. Nei vari ordini superiori *non si introducono numeri che non possano essere determinati con successioni fondamentali del primo ordine* (e questa è l'espressione cantoriana della «completezza» del campo reale); così facendo però si tien conto a parere di Cantor della «diversa forma e contenuto concettuale dell'esser dato» di un numero, sicché si avrebbe così a livello definitorio un'adeguazione della «generazione» di un particolare numero col suo modo «naturale» di presentarsi nella pratica del calcolo analitico.

2.3 *Definizione di Dedekind*

Dedekind parte dalla considerazione di *tutti* i numeri razionali, ossia considera l'intero insieme Q che intende come un dominio numerico *ordinato* di cui vengono assunte come rilevanti le seguenti caratteristiche: 1) si possono definire fra i suoi elementi le usuali relazioni di «uguale», «maggiore» e «minore» che godono della proprietà transitiva (ossia, per a, b, c, razionali: se a è maggiore [minore] di b e b è maggiore [minore] di c, allora a è maggiore [minore] di c); 2) l'*ordine* così determinato in Q è *denso* ossia dati due qualunque elementi a e b di Q diversi fra loro esiste sempre (almeno) un elemento c che «giace» fra di essi ossia che è maggiore del minore e minore del maggiore; 3) quella che

possiamo chiamare la proprietà di «sezionabilità»: fissato un numero razionale *a* qualsiasi, tutti i numeri di Q risultano divisi in due classi A_1 e A_2 entrambe infinite e tali che A_1 contiene tutti i razionali minori o uguali ad *a*, mentre A_2 contiene tutti i razionali maggiori di *a*. Risulta immediatamente che ogni numero della prima classe è minore di ogni numero della seconda; *a* stesso può essere indifferentemente assegnato alla prima classe come massimo o alla seconda classe come minimo (variando in modo ovvio la definizione).

Ora, nel saggio sugli irrazionali e la continuità, Dedekind effettua un confronto fra l'insieme Q considerato come sistema ordinato caratterizzato dalle proprietà precedenti e la retta *L* intesa come insieme ordinato di punti; individua quindi quella che a suo parere è l'«essenza» della continuità su *L* e propone poi la sua definizione deglii irrazionali che vede come un «completamento» dell'insieme Q che ponga questo insieme in condizioni di «rendere» questa ulteriore proprietà della retta. Più avanti ci occuperemo di questo; per ora limitiamoci a riferire la definizione dedekindiana di numero reale.

Le premesse per tale definizione vengono individuate da Dedekind nel procedimento «inverso» a quello che ci permette di constatare la proprietà di sezionabilità sopra vista: con questa si stabilisce che *ogni* numero razionale *a* produce una «sezione» nel senso sopra chiarito, ossia una ripartizione di tutti i numeri razionali in due classi opportune A_1 e A_2 a ciascuna delle quali indifferentemente il numero *a* può essere pensato appartenere. Si può esprimere in modo diverso la cosa dicendo che ogni numero razionale determina una sezione (A_1, A_2) tale che o la sua prima classe contiene un massimo o la sua seconda classe un minimo. Orbene, operando appunto il procedimento inverso, partendo cioè da partizioni già effettuate, è chiaro che in linea di principio possono presentarsi quattro possibilità nel senso che si può ottenere una partizione dei razionali in due classi A_1 e A_2 tali che: 1) A_1 ha un massimo e A_2 un minimo; 2) A_1 non ha massimo e A_2 non ha minimo; 3) A_1 ha un massimo e A_2 non ha un minimo; 4) A_1 non ha massimo e A_2 ha un minimo.

Ora, per la proprietà di densità prima riscontrata nell'ordine assunto dei razionali, è facile convincersi che non può aversi il caso 1). Si trovano invece facilmente esempi per i casi rimanenti; il

caso 3) si ha ad esempio ponendo in A_1 tutti i razionali $\leq 3/4$ e in A_2 i razionali rimanenti; il caso 4) ponendo ad esempio in A_1 tutti i razionali $< 5/6$ e in A_2 i rimanenti. Un esempio infine del caso 2) si ha ponendo, mettiamo, in A_1 tutti i razionali il cui quadrato sia minore di 2 e in A_2 i razionali rimanenti. Si nota subito che anche in quest'ultimo caso sono interessati *tutti* i numeri razionali e che le due classi risultanti da tale partizione sono infinite, sicché si ottiene una sezione dedekindiana. È altrettanto facile convincersi che in questo caso *non* esiste un numero razionale che produca questa sezione, ossia che funga da «elemento di separazione» fra le due classi (o altrimenti detto che sia o massimo della prima classe o minimo della seconda).

Orbene, afferma Dedekind, «ogniqualvolta abbiamo una sezione (A_1, A_2) non prodotta da alcun numero razionale, noi creiamo un nuovo numero, un numero *irrazionale*, che riguardiamo come completamente definito da questa sezione. Da ora in avanti quindi a ogni sezione definita corrisponde un numero definito razionale o irrazionale e noi riguardiamo due numeri come *diversi* o *non uguali* sempre e soltanto allora che essi corrispondano a sezioni essenzialmente diverse».

Riassumendo, i numeri irrazionali vengono ottenuti per Dedekind mendiante un «libero atto creativo dello spirito umano» e vengono quindi «associati» a particolari sezioni sull'insieme dei numeri razionali, precisamente a quelle sezioni non determinate da alcun numero razionale. Naturalmente si possono definire per le sezioni le usuali relazioni e operazioni fra numeri; la cosa è un po' laboriosa e delicata nel caso per così dire «ibrido», in cui siano implicate una «sezione razionale» e una «sezione irrazionale», ma non presenta difficoltà di principio.

Cantor sollevò alcune obiezioni alla definizione di Dedekind (come del resto a quella di Weierstrass) e questi a sua volta mise in evidenza quelle che a suo parere erano delle imperfezioni nella teoria cantoriana; si tratta tuttavia di obiezioni marginali, perlopiù di ordine tecnico, e che si manifestavano fra due autori ognuno dei quali riconosceva di buon grado la perfetta rigorosità dei metodi escogitati dall'altro (fra Dedekind e Cantor si era tra l'altro stabilita una salda amicizia personale). Possiamo quindi tralasciare di considerare queste obiezioni e osservare, con Dedekind, che in ogni caso «si prende una sola base comune, sulla quale ci si

deve essere accordati, l'aritmetica dei numeri razionali» e concludere, con Jean Cavaillès, che le tre teorie sopra esposte si differenziano, in realtà, soltanto per la diversa presentazione: «Dedekind procede in modo più descrittivo caratterizzando l'insieme dei numeri reali attraverso le sue proprietà, Weierstrass analizza il concetto di numero per trovarvi motivo e possibilità dell'estensione cercata, infine Heine e Cantor si preoccupano soprattutto dei procedimenti di calcolo». Ciò comporta che la definizione di Dedekind si fonda sulla pura struttura d'ordine dei razionali, mentre quella di Cantor e Weierstrass coinvolge anche la struttura algebrica (somme, prodotti, ecc.).

Ben più pregnanti, perché di natura generale, le obiezioni sollevate contro tutte e tre le teorie proposte (anche se in effetti ci si riduceva poi ad esaminare solo quella cantoriana e quella dedekindiana, come faremo anche noi) da Gottlob Frege nel secondo volume dei *Grundgesetze der Arithmetik* (*Principi dell'aritmetica*, 1903) e da Bertrand Russell nei *Principles of mathematics* (*Principi della matematica*, 1903). Tali critiche, pur indipendenti fra loro, colpiscono entrambe quello che è effettivamente un momento delicato delle teorie sopra esposte, e precisamente la questione della *dimostrazione dell'esistenza* dei numeri irrazionali in esse considerati. Viene così messo a fuoco un problema che sarà centrale in tutta la filosofia della matematica a partire proprio dal periodo che stiamo qui considerando e rispetto al quale si avranno posizioni assai differenziate: alludiamo al problema dell'*esistenza degli enti matematici* come costrutti insiemistici.

Si sarà notato che dopo aver posto le rispettive premesse definitorie i vari autori «associavano» (a particolari insiemi di razionali; o) alle successioni fondamentali o alle sezioni sui razionali, un nuovo «numero» la cui esistenza di fatto non veniva *dimostrata* in alcun modo; per essere più esatti tale esistenza veniva data per scontata, in modo decisamente confuso e non giustificato, sulla base di un assioma di continuità di cui parleremo più avanti. Schematizzando le cose – in vista anche della discussione che segue – possiamo quindi dire che se da questi autori veniva evitata la vecchia circolarità connessa con la definizione del numero come rapporto di grandezze geometriche, le rispettive teorie lasciavano sostanzialmente in sospeso due questioni di interesse fondamentale: 1) il problema se *esiste* un «oggetto» associabile a ognuno degli

elementi caratteristici delle varie definizioni, vale a dire che le definizioni non fossero «vuote»; 2) il problema se tale oggetto, una volta riconosciutane l'esistenza, si può considerare come un *numero* e se ha senso considerare la *totalità* dei numeri così ottenuti come in qualche modo data.

La difficoltà – rispetto alla questione analoga riguardante i numeri naturali – è che in questo caso i «numeri» che le definizioni individuavano sono oggetti *infiniti*, successioni o sezioni che vengono trattati come oggetti in sé, dati in atto, come oggetti compiuti e non come entità potenziali. È chiaro che a questo riguardo intervengono le convinzioni «filosofiche» dei vari autori. Cantor ha una decisa concezione realistica circa l'esistenza degli enti matematici: a partire da classi di oggetti esistenti, fornendo una definizione corretta (ossia, per lui, non obiettabile ma soprattutto *matematicamente* feconda) si individuano *ipso facto* enti della cui esistenza e realtà non c'è da dubitare. Con tale procedimento definitorio noi non facciamo altro che *descrivere* oggetti esistenti indipendentemente da ogni nostra considerazione linguistica o no, da ogni nostra esperienza, sensibile o no. Per Dedekind le cose stavano in modo completamente differente. Anche se egli si poneva il problema dell'esistenza degli irrazionali e lo risolveva con un «libero atto creativo della mente umana», aveva una visione che vogliamo chiamare funzionalistica, strutturale, della matematica. Come vedremo meglio più avanti, le caratteristiche con le quali Dedekind ritiene di individuare l'insieme dei razionali nella sua definizione sono in effetti (anche se nel '72 ancora implicitamente) di natura astratta, qualificano *ogni* sistema linearmente ordinato, denso e senza estremi come Q e non fanno riferimento a proprietà *intrinseche* dei razionali. È questa la caratteristica dell'approccio di Dedekind. La garanzia dell'esistenza di irrazionali è allora per lui data semplicemente dal fatto che su *ognuno* di tali *sistemi* si possono operare delle sezioni (infinite sezioni); l'attività creatrice della mente umana serve appunto a «generare» proprio *numeri* irrazionali, ossia specifici oggetti che intrattengono certe *relazioni* con *numeri* già noti e per i quali è possibile definire certe *operazioni* fra loro e con *numeri* già noti. Dedekind non si riferisce così a una data ontologia realisticamente intesa, ma fornisce un metodo generale per ottenere – dato un ordine denso – un ordine che lo completa nel senso sopra illustrato. La natura specifica degli enti che realizzano il

completamento è indifferente: ciò che interessa a Dedekind sono i rapporti strutturali più che la costruzione esplicita.

Al di là della loro portata diretta, le critiche di Frege e di Russell servono a mettere in luce il salto di consapevolezza metodologica che in un ventennio circa si determinerà nella ricerca matematica. La critica di Frege, almeno nei limiti in cui essa è condotta, colpisce in particolare il punto 1). Anche Frege è un realista convinto (e non a caso le obiezioni a Cantor si limitano sostanzialmente a evidenziare deficienze di tipo «logico») ma ritiene che una volta posta una definizione corretta non vi sia convinzione personale né potenza creatrice alcuna che possa garantire che essa non è «vuota»: è necessario *dimostrare* l'esistenza degli enti postulati dalla definizione stessa. Per quanto riguarda il secondo punto, invece, Frege non accetta il discorso dedekindiano (che qualifica *tout court* come formalista) semplicemente considerandolo superfluo: non si tratta di trovare isomorfie strutturali, bensì soltanto di presentare «oggetti» che godano di date proprietà. Comunque Frege, nell'opera citata, dopo una *pars destruens* estremamente dettagliata nella quale sottopone a critiche stringenti (anche se talora prolisse e venate di risentimento personale) le teorie presentate (con l'intento generale di dimostrare l'insostenibilità delle teorie formali dell'aritmetica), propone anche una propria soluzione al problema. Egli ribadisce la possibilità di definire il numero in generale come rapporto di «grandezze» e indica in effetti in modo ipotetico la natura di tali «grandezze» e il tipo di proprietà che certe classi di tali grandezze debbono possedere per rendere possibile una definizione logicamente corretta e *non vuota* (il che gli risulterebbe addirittura per costruzione) dei numeri reali. Il suo metodo avrebbe inoltre il non indifferente vantaggio di consentire una definizione *diretta* dei reali, senza cioè richiedere alcun appello all'«aritmetica dei razionali». Il fatto è che egli tronca la sua ricerca proprio lasciando aperto il compito di *dimostrare l'esistenza* di almeno una classe di «oggetti» (che Frege identificava con particolari relazioni) che giustifichino tale definizione. In altri termini, si può rivolgere contro la teoria abbozzata da Frege, paradossalmente, la stessa critica da cui egli aveva preso le mosse per «emendare» le teorie esaminate. Vanno tuttavia notate la chiarezza e la consapevolezza critica con le quali egli riesce a cogliere le deficienze logiche nelle definizioni che prende in esame, nonché

la franchezza con la quale denuncia l'insufficienza della propria costruzione.

Russell rileva le stesse difficoltà nelle teorie correnti degli irrazionali, e mostra come le tre teorie in questione siano sostanzialmente equivalenti su un piano puramente formale. Per quanto riguarda i punti 1) e 2) precedenti egli risolve il primo prospettando una teoria dei «segmenti» (da *non* intendersi in senso geometrico) di razionali con la quale in sostanza presenta un sistema la cui struttura è *isomorfa* (diremmo noi oggi) a quella richiesta per il campo numerico reale (in particolare per quanto riguarda la «continuità»); opera quindi una opportuna identificazione (che appunto mancava esplicitamente in Dedekind e Cantor) degli elementi di questo sistema di segmenti con i numeri razionali e con i «numeri» irrazionali. Per quanto invece riguarda il punto 2) la sua risposta è decisamente *negativa*: un numero reale *è* un particolare «segmento» e quindi non possiede in alcun modo, proprio in quanto oggetto, quelle caratteristiche «intuitive» che siamo soliti attribuire a un numero. Le due questioni sono quindi risolte nella stessa impostazione russelliana, operando su particolari sottoinsiemi ordinati dell'insieme dei numeri razionali (i «segmenti», appunto). Ciò permette a Russell di evitare l'*assioma di continuità* (si veda più avanti) perché «se esistono numeri razionali, devono esistere segmenti di razionali». Inoltre, col suo modo di affrontare il problema, Russell ritiene del tutto superata, come complicazione niente affatto necessaria, la questione di «creare» o comunque introdurre i *numeri* irrazionali «... giacché se i segmenti compiono tutto ciò che è richiesto dagli irrazionali, appare superfluo introdurre una nuova serie parallela di entità [gli irrazionali, appunto] con precisamente le stesse proprietà matematiche».

Ci sembra superfluo in questa sede scendere in ulteriori particolari tecnici. Vogliamo solo far notare come questioni del tipo sopra esposto vengano oggi di norma affrontate e «risolte» col metodo assiomatico risalente in sostanza a Hilbert (si veda il paragrafo 5) che comprende da un punto di vista più astratto le varie definizioni particolari proposte (giusta l'osservazione di Russell sulla loro sostanziale equivalenza formale), elimina in certo senso alla base la questione circa la *natura* geometrica, aritmetica, o altro degli enti che definisce implicitamente, e rimanda le questioni esisten-

ziali a un discorso semantico più generale, che porta in primo piano l'universo dei concetti astratti in cui tutte le nozioni matematiche vanno ricostruite. Ma avremo occasione di tornare nel seguito su questo aspetto della questione parlando della teoria degli insiemi e della Grande Logica.

2.4 *Numeri reali e continuo*

Si può dire che l'ottenuta definizione del concetto di numero irrazionale costituisse una soluzione soddisfacente (almeno da un punto di vista puramente matematico e sul metro di «rigore» di quel momento) del problema di «aritmetizzare» l'Analisi, almeno in due sensi: in primo luogo perché la rendeva indipendente dalla geometria (in quanto momento intuitivo e vago della fondazione) e anzi ribaltava il senso della dipendenza, fornendo all'indagine geometrica un valido strumento, ora autonomamente fondato; in secondo luogo perché assicurava all'Analisi, ossia alla scienza del continuo per eccellenza, un ben definito campo numerico «continuo» che ne precisava in modo automatico il riferimento fondamentale. Alla connessione fra teoria dei numeri reali, infinità e continuità (geometrica) non è stato possibile non accennare ripetutamente nelle pagine precedenti anche se, per chiarezza, si è tentato di isolare da questo contesto globale i vari elementi di questo classico problema. Vogliamo ora occuparci in particolare del problema della continuità.

Russell afferma in proposito, nei *Principi*: «Ci troviamo di fronte al problema che generalmente venne considerato come fondamentale della filosofia della matematica, intendo dire il problema dell'infinità e della continuità... Fin dai tempi di Newton e di Leibniz si è cercata la natura dell'infinità e della continuità attraverso le discussioni sul cosiddetto calcolo infinitesimale. Ma è stato dimostrato che questo calcolo non riguarda affatto, in realtà, l'infinitesimale, e che un grande e importantissimo ramo della matematica è logicamente antecedente a esso. Il problema della continuità è stato, inoltre, in gran parte separato da quello dell'infinità. Si supponeva un tempo, e qui sta la vera forza della filosofia della matematica di Kant, che la continuità avesse un riferimento essenziale allo spazio e al tempo, e che il calcolo infinitesimale (come sugge-

risce la parola *flussione*) presupponesse in qualche modo la nozione di moto o almeno di cambiamento. Secondo quest'ipotesi la filosofia dello spazio e del tempo precedeva quella della continuità... Tutto ciò è mutato per opera dei matematici moderni. Ciò che si chiama l'aritmetizzazione della matematica ha fatto vedere che tutti i problemi presentati, a questo riguardo, dallo spazio e dal tempo, sono già presenti nell'aritmetica pura.» Sicché risulta ora possibile «... dare una definizione generale di continuità, senza fare appello a quella massa di pregiudizi non analizzati che i kantiani chiamano "intuizione"; e... vedremo che lo spazio e il tempo non implicano alcuna altra continuità».

Per quanto riguarda più particolarmente quest'ultima nozione, notato che essa viene in generale considerata come non suscettibile di analisi, e assunta come primitiva, Russell osserva ancora che «si sono affermate molte cose su di essa, compreso il detto di Hegel, che ogni cosa discreta è anche continua e viceversa. Questa osservazione, che costituisce un esempio della solita abitudine di Hegel di associare gli opposti, venne pedissequamente ripetuta da tutti i suoi seguaci. Ma su ciò che essi intendevano per continuità e discretezza, essi hanno mantenuto un silenzio altrettanto continuo e discreto; una sola cosa era evidente: che qualunque cosa essi intendessero, questa non aveva importanza per la matematica, né per la filosofia del tempo e dello spazio».

Che cos'è allora il continuo e come l'analisi aritmetica si può sostituire all'intuizione geometrica? Nelle pagine che seguono schematizzeremo il problema e ci atterremo principalmente al corso delle argomentazioni di Dedekind nel suo saggio del 1872. Supponiamo quindi di possedere intuitivamente la nozione di continuità (geometrica) e di concretizzarla per comodità nella retta L intesa come insieme dei suoi punti: assumiamo cioè come noto intuitivamente il concetto di continuo (lineare) geometrico. Tentiamo quindi di caratterizzare tale continuo, di individuare cioè delle proprietà della retta che ragionevolmente si possano porre come caratteristiche e peculiari della proprietà da essa goduta di essere «continua». Un possibile modo di far ciò è il seguente: notiamo che la retta L può considerarsi come un sistema (totalmente) ordinato di punti, nel senso che due qualsiasi punti di L o coincidono oppure può dirsi che l'uno «precede» l'altro o viceversa che l'uno «segue» l'altro. Questa proprietà d'ordine è transitiva, ossia se A,

B, C sono punti distinti della retta e A precede B e B precede C allora anche A precede C. Altra proprietà che individuiamo facilmente è quella della densità, per cui cioè dati sulla retta due punti distinti A e B qualunque, esiste sempre (almeno) un punto C (e quindi infiniti) che «giace» fra di essi, ossia che «segue» A e che «precede» B. Si noti che, una volta pensato definito su L un sistema di misura, questa proprietà comporta subito che non esistono sulla retta misure minime: considerati infatti due punti distinti A e B, esisterà fra di essi un punto intermedio C la cui distanza da A e da B sarà minore della distanza fra A e B. Lo stesso ragionamento può ripetersi per le due coppie di punti A, C e C, B ora determinate e così via. Nella misura quindi in cui riteniamo la retta continua, avremo intanto ottenuto che se in un insieme ordinato esistono «distanze minime», tale insieme non potrà essere continuo. L'ultima proprietà cui vogliamo rifarci per caratterizzare la continuità è la seguente (sezionabilità): comunque si prenda un punto A sulla retta, questa viene divisa in due parti P_1 e P_2 entrambe infinite, e tali che ogni punto di P_1 precede ogni punto di P_2 (ogni punto di P_2 segue ogni punto di P_1); il punto A può essere indifferentemente assegnato come «ultimo» punto a P_1 o come «primo» punto a P_2.

Risulta chiaro che le proprietà con le quali abbiamo supposto di poter caratterizzare la continuità della retta L coincidono (salvo ovviamente il cambiamento di nomi) con le proprietà con le quali Dedekind aveva caratterizzato l'insieme Q dei numeri razionali. Ciò comporta immediatamente che se veramente avessimo con tali proprietà colto quella che Dedekind chiama l'«essenza» della continuità, dovrebbe risultare continuo anche l'insieme Q. Detto in altri termini – e a parte certe precisazioni – dovrebbe essere possibile associare ad *ogni* punto della retta senza esclusioni un numero razionale in modo per così dire che punti diversi ricevano «nomi numerici» razionali diversi. Eppure disponiamo, già di prove conclusive del fatto che ciò non è possibile: la scoperta dell'esistenza di segmenti incommensurabili è già sufficiente a mostrarci che la retta L è infinitamente più ricca di individui-punti di quanto Q non sia ricco di individui-numeri; vale a dire sappiamo già, grazie alla scoperta pitagorica dell'irrazionale, che non è possibile compiere l'associazione nominale di cui sopra si parlava: non esistono abbastanza «nomi numerici» (razionali) che ci per-

mettano di individuare punto per punto il continuo lineare geometrico. Ne discende subito che le caratterizzazioni sopra date per Q e per L sulla base delle quali esse risulterebbero per così dire «equivalenti» rispetto alla continuità, non hanno colto l'essenza della continuità della retta.

A questo punto intervengono le considerazioni «inverse» di Dedekind cui si accennava in riferimento alla sua definizione del concetto di numero irrazionale; infatti, se ora noi «... come desideriamo, vogliamo rendere compiutamente in modo aritmetico tutti i fenomeni sulla retta, il dominio dei numeri razionali è insufficiente e diventa assolutamente necessario che lo strumento Q costruito per mezzo della creazione dei numeri razionali sia essenzialmente migliorato mediante la creazione di nuovi numeri tali che il dominio dei numeri ottenga la stessa completezza, o come diremo, la stessa *continuità* della retta».

Ecco che allora il problema diventa quello di indicare una caratteristica precisa della continuità che possa farne una base per effettive deduzioni valide; e come il lettore avrà certamente compreso, tale caratteristica viene da Dedekind ritrovata appunto in una proprietà inversa a quella di sezionabilità. È vero infatti che ogni punto della retta produce una sezione nel senso sopra accennato; ma, viceversa, afferma Dedekind, «io trovo l'essenza della continuità nel converso, ossia nel seguente principio: "Se tutti i punti della retta si dividono in due classi tali che ogni punto della prima classe giace a sinistra di ogni punto della seconda classe, allora esiste uno e un solo punto che produce questa separazione di tutti i punti in due classi, questa suddivisione della retta in due parti"».

Si può ritenere, a prima vista, che questa caratterizzazione sia banale, ma a questo, dice Dedekind, «vorrei rispondere che... sono felice se ognuno trova il principio sopra esposto così ovvio e così in armonia con la sua propria idea di retta; poiché io infatti *non sono in grado di fornire alcuna dimostrazione della sua correttezza, né nessuno può far ciò. L'assunzione di questa proprietà della retta è null'altro che un assioma* in base al quale noi attribuiamo alla retta la sua continuità, per mezzo del quale troviamo la continuità nella retta» [corsivo nostro]. Espresso in termini di sezioni, quanto sopra significa semplicemente che sulla retta non è possibile eseguire sezioni del tipo 2) illustrato al numero precedente, o, tenendo conto del con-

tenuto intuitivo di tali sezioni, che sulla retta non esistono «lacune». Assumere quindi per la retta *L* la validità di questo principio comporta immediatamente mettere in risalto il diverso comportamento a questo riguardo dell'insieme Q dei razionali: in quest'ultimo si individuano infatti facilmente – come abbiamo già esemplificato – sezioni di tipo 2) ossia non determinate da alcun numero razionale. Orbene, proprio in «questa proprietà che non tutte le sezioni sono prodotte da numeri razionali consiste l'incompletezza o discontinuità del dominio Q di tutti i numeri razionali»; e risulta chiara ormai, d'altra parte, la motivazione dell'ampliamento *numerico* dell'insieme Q. Considerando ora non più l'insieme Q, ma l'insieme R ottenuto come riunione di tutti i numeri razionali e di tutti i numeri irrazionali così introdotti, Dedekind può dimostrare che R, oltre a godere delle tre proprietà sopra viste e caratteristiche per Q, gode anche appunto della proprietà di continuità nel senso del principio sopra specificato, vale a dire che ogni sezione effettuata in R possiede, in R, un elemento di separazione. Riprendendo qui per inciso il discorso sulla teoria dedekindiana degli irrazionali, è chiaro che una volta *postulata* l'esistenza di un elemento separatore (punto) per ogni sezione sulla retta, l'atto creativo non fa che consentire il passaggio dalla geometria all'aritmetica e non è particolarmente problematico per Dedekind, dopo quanto si è detto sulla sua impostazione di filosofia della matematica.

Va notato a questo punto che anche Cantor, già nella sua prima, succinta esposizione, tocca ovviamente il problema del rapporto fra insieme dei numeri reali e continuità della retta. In proposito afferma (la diversa terminologia non dovrebbe far sorgere difficoltà) che per «rendere completo il collegamento del dominio dei numeri reali sopra definito con la geometria sulla retta, va solo introdotto ancora un *assioma*, il quale semplicemente consiste in questo che reciprocamente anche a ogni numero reale corrisponde un punto determinato la cui coordinata (ascissa) è uguale a quel numero», aggiungendo che egli chiama «... assioma questa proposizione perché è nella sua natura di non essere in generale dimostrabile».

È estremamente importante notare come tanto Dedekind quanto Cantor colgano esplicitamente la natura assiomatica, postulazionale del concetto di continuità. Da una parte questo pone fine a

una ricerca antica di secoli sull'«essenza» della continuità tradizionalmente orientata verso una spiegazione riduzionistica, in termini di altri concetti, stabilendone la natura di «nozione primitiva» che viene spiegata tramite una caratterizzazione diretta e precisa, che illustra specifiche proprietà del concetto in questione; dall'altra questo riconoscimento concorre non poco, sia pure implicitamente, alla successiva revisione del metodo assiomatico che verrà puntualizzato nella sua forma moderna da Hilbert di lì a qualche decennio; più precisamente, mostra, da questo punto di vista, come esistessero già in questo periodo (e lo vedremo meglio nel paragrafo 5) germi precisi per tale cambiamento.

In questo ordine di idee, anticipando talune notizie per comodità del lettore e per completezza di esposizione, si possono fare alcune interessanti osservazioni sulla diversa portata dei due assiomi di continuità, quello dedekindiano (che assumiamo nella forma su esposta), e quello cantoriano, che in termini di una successiva enunciazione del suo autore può essere espresso come segue: «Data su una retta una successione di segmenti a_1, a_2,... a_n... aventi le seguenti proprietà:

1) ogni segmento contiene tutti i successivi;
2) prefissato comunque un segmento e piccolo a piacere esiste un segmento a_n della successione che ha lunghezza minore di e;

allora esiste uno e un solo punto interno a tutti i segmenti della successione.»

Questi due assiomi non sono infatti equivalenti; la ricerca successiva ha messo in luce che da quello di Dedekind (e dagli altri postulati abitualmente ammessi come fondamento della geometria) è derivabile una proposizione, nota come postulato di Eudosso-Archimede,[3] che *non* è invece derivabile da quello cantoriano. È immediatamente chiaro, pur senza scendere in particolari, che sarà quindi possibile, assumendo la continuità nel senso di Cantor, «costruire» geometrie che pur essendo continue *non* soddisfano il postulato archimedeo (le cosiddette geometrie non archimedee). Va infine detto che Hilbert, nella sua assiomatizzazione della geo-

[3] Il postulato di Eudosso-Archimede può essere espresso dicendo che dati sulla retta due segmenti distinti AB e CD, supposto ad esempio $AB > CD$, esiste sempre un n naturale tale che $nCD > AB$. In altri termini, dati due segmenti diversi, esiste sempre un multiplo del minore che supera il maggiore.

metria, esprime appunto la continuità mediante due proposizioni (assiomi) una delle quali è equivalente al postulato di Archimede, mentre l'altra, e la cosa dovrebbe risultare comprensibile al lettore dopo quanto detto, afferma proprio la non estendibilità, la completezza del sistema così ottenuto ove siano soddisfatti tutti gli altri assiomi.

Senza discutere in questa sede tutta una serie di altri interessantissimi problemi che questa nuova concezione del continuo pone (ad alcuni dei quali accenneremo in seguito) vogliamo qui ricordare due classiche questioni filosofiche in questo contesto, che furono poste in una luce completamente nuova dalle scoperte matematiche sopra delineate.

1) La prima è il problema di determinare quanto questa caratterizzazione matematica del continuo risulti indipendente dalla «materia» sulla quale è stata costruita, in particolare dall'assunzione dalla retta come modello di continuità. Il fatto importante (provato da Cantor) è che si può dimostrare che *ogni* altro tipo di continuo si può riportare a quello aritmetico, che giustifica così il suo ruolo nella ricostruzione del continuo temporale e spaziale. In connessione con le riflessioni conseguenti la scoperta delle geometrie non euclidee da una parte e con le idee di Riemann e Helmholtz dall'altra, viene sempre più imponendosi a partire dalla metà dell'ottocento una visione più articolata della continuità dello spazio e del suo ruolo. Afferma ad esempio Dedekind: «Ammesso che lo spazio abbia un'esistenza reale, *non* è per esso necessario essere continuo; molte delle sue proprietà rimarrebbero uguali anche se fosse discontinuo. E se noi sapessimo per certo che lo spazio non è continuo, nulla ci impedirebbe di colmare le sue lacune, rendendolo così continuo; questo colmare le lacune consisterebbe in una *creazione* di nuovi individui-punti e verrebbe effettuata secondo il principio di continuità sopra enunciato». E Cantor, da parte sua, ritiene che «*l'ipotesi della continuità dello spazio* non è null'altro che l'assunzione, in sé arbitraria, della corrispondenza biunivoca e completa fra il *continuo* tridimensionale *puramente aritmetico*... e lo spazio posto a fondamento del mondo dei fenomeni» [corsivi nostri].

2) Altro problema è quello del rapporto fra continuo così concepito e gli elementi che implicitamente si ritiene lo costituiscano. Le soluzioni proposte sono a tutt'oggi assai diverse e va menziona-

ta in particolare la posizione intuizionista (che avremo occasione di discutere più avanti in dettaglio) secondo la quale il continuo è il fatto originariamente dato dal quale poi si «estraggono» i «punti costituenti» e non viceversa.

Ma è nel rapporto continuità-infinità che si fa strada un'altra problematica che, da una sua peculiare angolatura porterà ad altre caratterizzazioni di tipo ordinale e di tipo cardinale del continuo stesso: che costruzioni si possano fare – analoghe alle sezioni dedekindiane – con insiemi qualunque? Concetti quali quelli di numero e di ordinamento si possono estendere a collezioni qualsiasi di oggetti? È affrontando questioni di questo tipo che si svilupperà la teoria degli insiemi di Cantor di cui ci occuperemo nel paragrafo 3.

2.5 *L'assiomatizzazione dell'aritmetica*

Nel contesto dell'esigenza di rigore e generalizzazione della matematica del XIX secolo, l'aritmetizzazione dell'Analisi si poneva quindi come un primo essenziale momento, conclusivo del grande processo di revisione critica teso ad assicurare alla matematica una fondazione autonoma e una posizione in certo senso privilegiata rispetto alle altre scienze. Felix Klein esprime molto bene questa concezione nel 1895: «In Euclide – e in generale negli antichi pensatori matematici – la geometria è, grazie ai suoi assiomi, la base rigorosa dell'aritmetica generale, che comprende anche l'irrazionale. L'aritmetica ha conservato questa condizione di vassallaggio rispetto alla geometria fino al XIX secolo. Ma in seguito le condizioni si sono completamente modificate; oggi è proprio l'aritmetica che ha ottenuto il predominio, come la vera disciplina fondamentale.»

Dopo che si era liberato da questo «vassallaggio» geometrico il concetto generalissimo di numero reale (e con esso l'Analisi), quello che con Weierstrass era divenuto il «problema di coscienza» della matematica ottocentesca – il problema cioè di una definizione rigorosa del concetto di numero – si specializza per così dire nella questione di dare una chiarificazione puramente «aritmetica» dello stesso concetto di *numero naturale*. Si specializza e non si «riduce» perché nella misura in cui l'aritmetica, questa «regina

della matematica»[4] poteva considerarsi come la struttura teorica fondamentale, risulta ben presto evidente che una giustificazione fondante del concetto di numero naturale, un chiarimento della natura delle leggi aritmetiche assumevano la portata generalissima di una più precisa collocazione della conoscenza e della stessa attività matematica nell'ambito di una visione filosofica globale della conoscenza.

In questo contesto, le posizioni dei vari studiosi cominciano a differenziarsi. Così, ad esempio, J. Stuart Mill, dopo che l'assolutezza del concetto di spazio e la stessa natura aprioristica della geometria erano state messe in crisi dalla scoperta delle geometrie non euclidee, tenta di inserire in un quadro empirista (di marca psicologistica) anche l'aritmetica (ossia, in definitiva, la matematica). Il numero è per lui una pura *rappresentazione* soggettiva dello spirito umano, e dietro a ogni numero c'è un *fatto*; le leggi aritmetiche non sono a suo parere che generalizzazioni empiriche, soggette quindi a variare al mutare delle contingenti condizioni del mondo esterno che le determinano. Da un punto di vista psicologistico prende le mosse anche Helmholtz, che in un volume del 1887 definisce l'aritmetica come «... un metodo costruito su fatti puramente psicologici, in base ai quali viene assegnata l'applicazione coerente di un sistema di simboli di portata illimitata e con illimitata possibilità di affinamento».[5] A partire poi dalla psicologia dell'atto del contare, Helmholtz ritiene di poter giungere all'enunciazione di un sistema completo di assiomi per l'aritmetica, nel tentativo di gettare un ponte fra aritmetica e esperienza (si pensi alle sue analoghe posizioni riguardo alla geometria). Nello stesso volume in cui appare il contributo di Helmholtz, Leopold Kronecker (1823-91) compendia nel suo famoso detto «Dio ha creato i numeri naturali, il resto è opera dell'uomo» quella che

[4] Era, come noto, l'appellativo riservato da Gauss alla teoria dei numeri: «La matematica è la regina delle scienze e la teoria dei numeri è la regina della matematica.»

[5] La citazione è tratta dal volume *Philosophische Aufsätze zu E. Zellers 50-jährigen Doktor jubileum* (*Memorie filosofiche per il 50° anniversario del dottorato di E. Zeller*), pubblicato a Lipsia nel 1887. Helmholtz vi contribuì con l'articolo *Zählen und Messen, erkenntnistheoretisch betrachtet* (*Contare e misurare, dal punto di vista della teoria della conoscenza*) e Kronecker con l'articolo *Über den Zahlbegriff* (*Sul concetto di numero*).

può essere considerata una prima «grossolana» manifestazione di un atteggiamento costruttivista e già dal 1879 Frege tenta ripetutamente di imporre negli ambienti scientifici tedeschi il suo programma logicista sui fondamenti della matematica.

Una posizione in certo senso intermedia fra quelle sopra delineate (e di alcune delle quali avremo occasione di riparlare nel corso delle nostre pagine) assume Dedekind nel saggio *Essenza e significato dei numeri* del 1888 (seconda edizione 1893); posizione che a noi qui particolarmente interessa non tanto e non solo per l'intrinseco contenuto epistemologico ma anche perché essa sfocia in una prima assiomatizzazione dell'aritmetica. Dedekind dichiara esplicitamente di essersi deciso a pubblicare il suo saggio dopo aver preso conoscenza dei contributi sopra citati di Helmholtz e di Kronecker; e altrettanto chiaramente afferma che se pure le sue idee sono «per molti aspetti simili» a quelle di tali autori, tuttavia «in sostanza sono da esse essenzialmente differenti».

Si vedrà che in certo senso si potrebbe infatti parlare di uno «psicologismo» di Dedekind, che si manifesterebbe nel rimando alle «leggi del pensiero» per la giustificazione del numero; in questo senso Dedekind è in effetti l'ultimo anello di una catena che lo unisce a Boole, probabilmente con la mediazione di Schröder, che egli cita spesso e per il quale dichiara la massima stima e ammirazione. Ma l'eventuale momento psicologico interviene in Dedekind, a nostro parere, come «appendice» terminale di un discorso fondante che in effetti si svolge su un piano puramente aritmetico (e quindi, per Dedekind, logico).

La ricerca di Dedekind non è diretta a mettere in evidenza gli *oggetti* ultimi della conoscenza aritmetica (i numeri cioè) bensì a scoprire i *processi* che eventualmente possono generarli, le *relazioni* che fra di essi possono stabilirsi ed il loro ruolo nella conoscenza. È sì – come vedremo – una precipua «capacità» della nostra mente a permetterci la costituzione della successione numerica, ma tale capacità è specificamente matematica, è la possibilità di *Abbildung*, di rappresentazione, che viene intesa in senso puramente «formale», e cioè come operante astrattamente e su «cose» lasciate di principio indeterminate. L'atto creativo che conclude tanto nel caso degli irrazionali quanto, come vedremo, nel caso dei naturali, questo procedere dedekindiano, ha quindi più lo scopo di *individuare* specifici oggetti, che non quello di *fondare* l'aritmetica,

trovare l'ambito di validità dei suoi teoremi. La fondazione è astratta e strutturale e a nostro parere l'atteggiamento di Dedekind anticipa la posizione hilbertiana; è indubbio tuttavia che un'interpretazione puramente «nominalistica» di quello che abbiamo altrove chiamato lo «strutturalismo» dedekindiano ebbe non poca responsabilità sulle manifestazioni prehilbertiane di un formalismo ingenuo, che riconosceva nel segno in quanto tale l'essenza del numero, e che eliminava così alla radice ogni questione esistenziale.

Nel suo saggio Dedekind assume come fondamentale il canone che tutto ciò che nella scienza è suscettibile di dimostrazione debba essere dimostrato nel senso che se ne deve stabilire il fondamento, le ragioni di validità. Ritiene da questo punto di vista insoddisfacente la situazione della «più semplice di tutte le scienze», dell'aritmetica, della quale parla come di una «branca della logica» nel senso che egli considera il concetto di numero come «interamente indipendente dalle nozioni o intuizioni di spazio e tempo», per considerarlo invece come un «risultato immediato delle leggi del pensiero». La risposta che Dedekind dà alle domande che costituiscono il titolo del proprio saggio è brevemente la seguente: «I numeri sono libere creazioni della mente umana; essi servono come mezzo per apprendere più facilmente e con maggior sottigliezza la differenza fra le cose». È il ruolo dei numeri quindi che conferisce ad essi il carattere logico, ed è in questo loro aspetto che sta la loro essenza.

In contrapposizione allo psicologismo helmholtziano, che finiva col trarre gli assiomi dell'aritmetica dalla pura forma intuitiva del tempo che è la base dell'atto del contare e ne costituisce la condizione di possibilità, Dedekind prosegue affermando: «È solo attraverso il *processo puramente logico* di costituzione della scienza dei numeri con la conseguente acquisizione del dominio continuo dei numeri che noi siamo accuratamente preparati a indagare le nostre nozioni di spazio e di tempo, portandole in relazione col dominio numerico creato dalla nostra mente. Se analizziamo più accuratamente ciò che facciamo contando un aggregato o un numero di cose, siamo condotti a considerare la *capacità della mente a collegare* cose a cose, a far *corrispondere* cose a cose, o a *rappresentare* una cosa per mezzo di un'altra, capacità senza la quale non sarebbe neppure possibile pensare. *Su quest'unica e quindi assolutamente indi-*

spensabile fondazione... dev'essere a mio parere basata l'intera scienza dei numeri» [corsivi nostri].

Già nel saggio del '72 Dedekind aveva affermato di riguardare l'aritmetica come «una conseguenza necessaria o almeno naturale, del più semplice atto aritmetico, quello del contare», precisando poi che il contare stesso non era a suo parere altro che «la successiva creazione della successione infinita di interi positivi nella quale ogni individuo è definito per mezzo di quello che lo precede immediatamente; la più semplice attività consiste nel passare da un individuo già formato al consecutivo nuovo da formarsi». Questa precisazione è importantissima, perché mette in luce come già allora Dedekind assegnasse una funzione centrale all'induzione matematica nella costituzione e individuazione della successione dei numeri naturali, evitando ogni discorso su unità, quantità e simili, e come quindi c'era da aspettarsi che una precisazione dell'induzione stessa sarebbe intervenuta a livello *definitorio* in una più precisa sistemazione della materia. Si comprende allora che nel saggio dell'88 Dedekind annoveri fra i risultati fondamentali che gli sembra di aver raggiunto, oltre a una precisa definizione dell'infinito (che riterrà punto essenziale di superiorità anche nei riguardi della teoria degli insiemi che Cantor veniva in quel periodo sviluppando), proprio la dimostrazione del fatto «che quella forma di argomento nota come induzione completa (ossia l'inferenza da n a $n + 1$) è effettivamente conclusiva... e che di conseguenza la definizione per induzione (o ricorsione) è determinata e coerente».

Vediamo ora come Dedekind sviluppa il suo discorso. Egli prende le mosse dalla considerazione di sistemi (insiemi, classi) arbitrari S di «cose» (col che egli intende «ogni oggetto del nostro pensiero») e definisce fra di essi le relazioni e operazioni insiemistiche elementari di inclusione, riunione e intersezione. Consideriamo ora due sistemi qualunque S e T; la fondamentale nozione di *Abbildung*, rappresentazione o trasformazione del sistema S nel sistema T, è intesa come una qualsiasi legge o prescrizione (oggi diremmo *funzione*) f in base alla quale a ogni determinato elemento s di S corrisponde un elemento $f(s)$ di T, detto trasformato o immagine di s. Se la f è iniettiva (ossia tale che a $s \neq s'$ di S corrispondono come immagine elementi $f(s)$ e $f(s')$ distinti di T) allora Dedekind chiama la trasformazione *simile* o *distinta*. Fra le varie

rappresentazioni f che possono aversi per un sistema S, Dedekind fissa ora l'attenzione su quelle che rappresentano il sistema in se stesso, vale a dire sono tali che per ogni $s \in S$ si abbia $f(s) \in S$.[6] Un sistema che goda di questa proprietà *rispetto a una data trasformazione f* viene detto una *catena* (rispetto a f).

Consideriamo ora un sistema S e una trasformazione f di S in se stesso. Preso un *qualunque* sottoinsieme A di S, si considerino tutte le catene (rispetto ovviamente a f) che contengono A (di tali catene ne esistono certamente: S stesso ad esempio è una di esse, evidentemente la «massima»). Si faccia ora l'intersezione[7] di tutte le catene che contengono A. Si dimostra che tale intersezione è ancora una catena: essa viene detta *catena di A* e indicata con A_0. A questo punto Dedekind può dimostrare *per catene in generale* quello che ora diventa il *teorema* di induzione completa, ossia la seguente proposizione: Per dimostrare che gli elementi di una catena A_0 godono di una certa proprietà P, è sufficiente dimostrare che

a) tutti gli elementi di A godono di P;

b) se $a \in A_0$ gode di P allora $f(a)$ gode di P.

È appunto questo risultato che costituisce la «base scientifica per quella forma di dimostrazione nota come induzione completa».

Dedekind introduce a questo punto la sua celebre definizione di sistema infinito (a proposito del quale sorsero sgradevoli questioni di priorità con Peirce), intendendo come tale un sistema S che sia rappresentabile *biunivocamente* su un suo sottoinsieme proprio S', vale a dire mediante una funzione iniettiva f che *esaurisce* S (sia cioè anche *suriettiva*, si veda la nota 17, pag. 309), e passa a definire quindi il concetto di sistema *semplicemente infinito*: un sistema S viene detto s.i. quando esiste una trasformazione iniettiva f di S in

[6] Il segno « $\in$ », detto segno di appartenenza, abbrevia dizioni quali «appartiene», «è elemento di», sicché un'espressione della forma « $a \in A$ » indica il fatto che l'elemento a appartiene all'insieme A (o, come anche si dice, che a è membro o elemento di A). $\in$ denota cioè la relazione di appartenenza di un elemento a un insieme.

[7] Data una famiglia F di insiemi, l'*intersezione* $\bigcap F$ è costituita da tutti e soli gli elementi che appartengono a *tutti* gli insiemi $X \in F$; la *riunione* $\bigcup F$ è data invece da tutti e soli gli elementi che appartengono a uno almeno degli $X \in F$. Nel caso F contenga i due soli elementi X e Y, scriveremo rispettivamente $X \cap Y$ e $X \cup Y$. Se poi la famiglia è indiciata dall'insieme I, cioè se $F = \{X_i\}_{i \in I}$, scriveremo rispettivamente $\bigcap_{i \in I} X_i$ e $\bigcup_{i \in I} X_i$.

se stesso tale che S possa esprimersi come catena di un suo elemento non contenuto in $f(S)$. Conveniamo con Dedekind di indicare col *simbolo* 1 tale elemento. Risulta allora dalla definizione e dalle proprietà delle catene che «l'essenza di un sistema semplicemente infinito S consiste nell'esistenza di una trasformazione f di S e di un elemento 1 che soddisfano le seguenti condizioni

α. $f(S) \subseteq S$[8]
β. $S = 1_0$
γ. L'elemento 1 non è contenuto in $f(S)$
δ. La trasformazione f è iniettiva».

È chiaro che un sistema semplicemente infinito può concepirsi come una successione *generata e ordinata* dalla trasformazione f; a partire infatti dall'elemento 1 di S, che per definizione non è immagine di alcun altro elemento di S, si otterrà successivamente

$$1, f(1), f(f(1)), f(f(f(1))),$$

e così via.

È proprio a questo punto che si coglie l'«essenza», è il caso di dirlo, dell'impostazione dedekindiana. Se infatti – afferma Dedekind – «nella considerazione di un sistema semplicemente infinito S ordinato da una trasformazione f trascuriamo completamente il carattere peculiare degli elementi, ritenendo semplicemente la loro distinguibilità e *tenendo conto soltanto delle relazioni fra essi determinate dalla trasformazione f*, allora questi elementi sono chiamati *numeri naturali* o *numeri ordinali*, o semplicemente *numeri*... In riferimento *a questo liberare gli elementi da ogni loro contenuto* (astrazione) *siamo giustificati nel chiamare i numeri una libera creazione della mente umana*» [corsivi nostri].

Ci sembra che queste parole siano sufficienti a chiarire definitivamente la marginalità, la sostanziale «estraneità» del processo creativo in quanto inteso come momento «psicologico» nell'ambi-

[8] Il segno «$\subseteq$», detto segno di inclusione, abbrevia dizioni quali «è incluso», «è contenuto» e simili, sicché un'espressione della forma $A \subseteq B$ indica il fatto che l'insieme A è incluso nell'insieme B (o, come di norma si dice, A è un *sottoinsieme* di B). È chiaro che la relazione di inclusione simbolizzata da «$\subseteq$» sussiste fra due insiemi A e B allora e solo allora che tutti gli elementi di A sono elementi anche di B. È ovvia la sostanziale differenza fra la relazione di inclusione e quella di appartenenza (si veda la nota 6 precedente); si osservi tuttavia che Dedekind usa lo stesso simbolo per entrambe.

to globale dell'impostazione dedekindiana. Un momento soggettivo è semmai da cogliere, a nostro parere, nella concezione ancora ingenua e intuitiva che Dedekind sembra avere dell'astrazione. È una concezione che resterà ancora a lungo radicata nel discorso «fondazionale» malgrado che già nel 1884 nelle sue *Grundlagen*, Frege ne avesse data una soddisfacente giustificazione in termini puramente logici (si veda il paragrafo 4); per sua stessa ammissione, Dedekind non conosceva ancora nell'88 il volume di Frege che pure come vedremo (e come del resto Dedekind stesso riconosce nella seconda edizione del suo saggio) si muoveva per certi versi, in particolare rispetto all'induzione, «in stretta connessione» col suo discorso.

Ma a parte questo, la concezione strutturalista di Dedekind viene manifestata in tutta la sua portata quando egli afferma che «le relazioni o leggi che sono denotate interamente dalle condizioni α β, γ, δ, e quindi sono sempre le stesse in tutti i sistemi ordinati semplicemente infiniti, *qualsiasi nome si dia agli elementi individuali*, costituiscono l'oggetto immediato della *scienza dei numeri o aritmetica*». In altri termini abbiamo definito una *struttura* relazionale in termini puramente matematici (logici): in questo sta l'essenza dell'aritmetica. Il successivo atto creativo, giustificato proprio dall'astrattezza e quindi generalità della struttura individuata, non fa al limite che dare un corpo, o se si preferisce un nome proprio definito, agli elementi di un qualche insieme particolare fra i quali sussistano quelle relazioni. Se allora col *simbolo* 1 intendiamo denotare proprio il *numero* 1, la successione N dei numeri naturali in senso intuitivo verrà per definizione ad essere quella particolare catena numerica (che Dedekind dimostra essere unica) che contiene (come elemento base, ossia non immagine di nessun altro) il numero 1.

Non ci sembra necessario discutere, in questa sede, l'articolazione del saggio di Dedekind, e trascuriamo quindi alcuni importanti argomenti e risultati che vi vengono trattati o raggiunti;[9] va

[9] Ad esempio la prima trattazione generale in termini insiemistici delle definizioni per ricorsione. Dedekind dimostra un teorema generale (il 126, paragrafo IX del suo saggio) sulle definizioni per induzione (come egli le chiama) e quindi lo applica alla definizione di somma (paragrafo XI), di prodotto (paragrafo XII), elevamento a potenza (o involuzione, nella sua terminologia, paragrafo XIII) fra naturali. È ancora degno di menzione il fatto che nel paragrafo XIV Dedekind conduce il primo confronto fra due diverse definizioni di infinito: quella nel suo

però messa in evidenza, per concludere il discorso, la dimostrazione (che Dedekind conduce nel paragrafo X del suo saggio) del fatto che due qualunque sistemi semplicemente infiniti M e M^* sono fra loro *isomorfi* relativamente alle rispettive trasformazioni Φ e Φ^* e ai rispettivi elementi base 1 e 1*: esiste cioè fra di essi una corrispondenza biunivoca f tale che 1 corrisponde a 1* (ossia $1 = f(1^*)$ e viceversa) e se $a \in M$ corrisponde ad $a^* \in M^*$ ($a = f(a^*)$ e viceversa) allora anche $\Phi(a) \in M$ corrisponde a $\Phi^*(a^*) \in M^*$ ($\Phi(a) = f(\Phi^*(a^*)$ e viceversa). Questa *dimostrazione* è molto importante e anzi *essenziale* ai fini di Dedekind: solo grazie ad essa egli può infatti affermare di aver *completamente determinato a livello definitorio* la struttura astratta degli insiemi semplicemente infiniti.

In questa dimostrazione ha una parte decisiva e fondamentale proprio l'induzione o, in termini dedekindiani, il concetto di *catena di*. Considerato l'elemento base 1, risulta infatti, grazie alla definizione di 1_0, che in essa sono contenuti *tutti* gli elementi raggiungibili con l'applicazione di Φ (perché, in quanto intersezione, 1_0 è la *parte comune* di tutte le catene che contengono 1) e d'altra parte anche *solo* quegli elementi (in quanto, ancora, come intersezione, 1_0 è la *minima* parte comune). Applicato alla catena N dei numeri naturali: in essa sono compresi tutti e soli gli elementi raggiungibili dal *numero* 1 con un numero *finito* di applicazioni dell'operazione di *successore* (che è la Φ di questo caso particolare). In altri termini, il processo d'induzione permette a Dedekind di escludere dalla successione dei numeri naturali ogni altro eventuale elemento «estraneo»: in questo senso Dedekind pensa di aver *caratterizzato univocamente*, da un punto di vista strutturale, l'insieme N.[10]

La determinazione assiomatica dell'aritmetica che, pur operan-

senso, e quella nel senso intuitivo di «non finito». Per dimostrare l'equivalenza di tali definizioni Dedekind è costretto a fare un inconsapevole appello all'assioma di scelta, per il quale si confronti il paragrafo 3.

[10] Ove si convenga di considerare la sistemazione di Dedekind come un'assiomatizzazione *esplicita* (si veda del resto quanto detto più avanti sul sistema di assiomi di Peano) in termini moderni si può dire che Dedekind ha dimostrato la *categoricità* del suo sistema d'assiomi, ha fatto cioè vedere che due qualunque *modelli* di quegli assiomi sono fra loro isomorfi. Da questo punto di vista, la correttezza della sua dimostrazione può essere assunta solo sulla base di tutta una serie di precisazioni e distinzioni che la critica moderna ha messo in evidenza a livello espressivo (linguistico) e deduttivo (logico inferenziale).

te come abbiamo creduto di mostrare, resta tuttavia solo implicita nella sistemazione dedekindiana, venne esplicitamente proposta per la prima volta nel 1889 dal matematico e logico italiano Giuseppe Peano (1858-1932) negli *Arithmetices principia nova methodo exposita* (*Principi dell'aritmetica esposti con metodo nuovo*). Il volume è scritto in latino ma in effetti del latino Peano si serve solo nella prefazione e per proporre un sistema di *Logicae notationes* che costituiscono non certamente il primo (al solito bisogna anche in questo caso far posto a Frege) ma certamente il più fortunato e indovinato fra i simbolismi sistematici escogitati in questo campo. Noi comunque, per comodità del lettore, enunceremo gli assiomi in linguaggio comune.

Va subito detto che l'assiomatizzazione di Peano è nel contempo un progresso e un regresso rispetto alla sistemazione dedekindiana (di cui, come vedremo, è sostanzialmente una *traduzione* in termini esplicitamente assiomatici). È un progresso appunto perché esplicita e perché simbolica; ancora, perché in essa viene chiaramente enunciato come *assioma* il principio di induzione completa. È invece un regresso perché è raffrontabile con un'assiomatica euclidea di tipo classico piuttosto che con un'assiomatica moderna di tipo hilbertiano di cui, come abbiamo detto, le conclusioni di Dedekind possono in certo senso essere considerate una realizzazione *ante litteram* per l'aritmetica.[11] Ci sembra in altri termini di poter dire che mentre nel caso di Dedekind siamo di fronte a un tentativo di *fondazione* dell'aritmetica, nel caso di Peano siamo di fronte invece a un tentativo, anche se indubbiamente significativo, di *sistemazione* della teoria dei numeri.

Ma consideriamo ora il sistema assiomatico di Peano. Esso è costituito da nove assiomi, quattro dei quali relativi all'identità (i primi tre descrivono l'identità come una relazione riflessiva, simmetrica e transitiva; il quarto afferma che se due oggetti sono uguali e uno di essi è un numero allora anche l'altro è un numero) mentre i rimanenti cinque sono «specifici» per l'aritmetica. Que-

[11] Ma si tenga presente che nel tomo secondo del *Formulario*, del 1898, Peano afferma esplicitamente: «La composizione del mio lavoro a.1889 fu ancora indipendente dallo scritto menzionato del Dedekind [*Essenza e significato*...]; prima della stampa, ebbi la prova morale dell'indipendenza delle proposizioni primitive da cui io partivo, nella loro coincidenza sostanziale colle definizioni del Dedekind».

sti sono (a destra, tra parentesi, riportiamo la numerazione originale):

1) 1 è un numero (1);
2) il successore di un numero è un numero (6);
3) se due numeri sono uguali allora hanno successori uguali (7);
4) 1 non è il successore di alcun numero (8);
5) se k è una classe tale che 1 appartiene a k e inoltre per ogni numero x se x appartiene a k allora anche $x + 1$ appartiene a k, allora k contiene la classe N (9).

A questi assiomi Peano premette quattro *explicationes* con le quali fissa, predetermina, il significato dei quattro segni che compaiono nella versione simbolica degli assiomi stessi (il lettore non avrà difficoltà a «immaginare» dove quei segni intervengono): «1 significa unità; $a+1$ significa successore; N significa numero (intero positivo); = significa è uguale (identità)». Non sarà difficile per il lettore rendersi conto del fatto che gli assiomi precedenti altro non sono che la «traduzione» delle quattro proposizioni dedekindiane sui sistemi semplicemente infiniti; e infatti si vede immediatamente che l'assioma 2) corrisponde alla proposizione α di Dedekind; l'assioma 3) alla δ; l'assioma 4) alla γ; infine gli assiomi 1) e 5) corrispondono alla proposizione β. Si noti tuttavia che mentre per Dedekind queste proposizioni determinavano un'intera classe di strutture possibili, per Peano, viste le *explicationes* di cui sopra, il sistema assiomatico veniva vincolato a caratterizzare in modo univoco, ontologico, *il* sistema dei *numeri naturali*. Così, mentre Dedekind si poneva il problema di dare una *dimostrazione* di categoricità, Peano assumeva intuitivamente come data questa proprietà e semplicemente l'*affermava* con i suoi assiomi.

La cosa può essere meglio compresa se si pensa alle considerazioni con cui Russell mostrò che il sistema d'assiomi di Peano non caratterizzava *intrinsecamente* l'insieme dei numeri naturali, ma sostanzialmente qualunque progressione: si pensi semplicemente di sostituire di volta in volta alle *explicationes* peaniane le seguenti: 1 significhi il numero 100 (o qualunque altro numero naturale) e gli altri simboli conservino l'usuale significato; 1 significhi il numero 2, successore significhi $a + 2$, N significhi insieme di numeri pari; e così si potrebbe continuare. Orbene, queste determinazioni dei significati dei segni, e i sistemi relativi cui esse danno luogo, mettono in crisi la pretesa caratterizzazione di Peano; ma non toccano

da questo punto di vista quella puramente strutturale di Dedekind, dal momento che i sistemi così ottenuti sono ancora tutti isomorfi fra loro. Per poter parlare della «non categoricità» dell'aritmetica in questo senso è necessario far ricorso a considerazioni di tipo più «profondo» che coinvolgono la determinazione precisa dei sottoinsiemi di N cui si suppone faccia riferimento lo schema d'induzione, e quindi la scelta del linguaggio in esame. Ma di questo parleremo nel capitolo V.

Va ancora tenuto conto del fatto che Peano, nell'affrontare questi problemi manteneva consapevolmente un atteggiamento per così dire «afilosofico»; il suo scopo era semplicemente quello di escogitare un simbolismo agile ed espressivo, capace di dare una presentazione chiara ed efficace di teorie e risultati matematici. In questo senso egli pubblica nel 1894 le *Notations de logique mathématique* (*Notazioni di logica matematica*) e lo stesso scopo persegue nelle cinque edizioni, dal 1895 al 1908, del *Formulaire de mathématique* (*Formulario matematico*) che conduce con l'aiuto di tutta una schiera di valenti collaboratori. L'ultima edizione del *Formulario* è in *latino sine flexione*, la lingua che Peano aveva escogitato come linguaggio internazionale. È interessante riportare da questa le parole con le quali egli introduce il proprio sistema assiomatico (che dopo la prima versione subisce alcune modifiche non sostanziali; ad eccezione di quella, suggerita quasi certamente dalle analisi di Frege, di far cominciare N con lo 0 anziché con l'1): «Quaestione si nos pote defini N_0 significa si nos pote scribe aequalitate de forma

N_0 = expressione composito per signos noto $\cup$, $\cap$, $-$, ... ι, quod non es facile.

«Ergo nos sume tres idea N_0, 0, + ut idea primitivo, per que nos defini omni symbolo de Arithmetica.

«Nos determina valore de symbolo non definito N_0, 0, + per systema de propositio primitivo sequente» a cui seguono gli assiomi sopra detti.

Riprenderemo nell'ultimo paragrafo il discorso sugli sviluppi che – con Peano e la sua scuola – la ricerca fondazionale ebbe in Italia. Volgiamoci ora ai due fatti più significativi dal nostro punto di vista e che maturano a partire dagli anni settanta del secolo scorso: l'analisi cantoriana del concetto di insieme e la nuova angolazione in cui Frege pone la logica e, con essa, l'aritmetica.

3. LA TEORIA DEGLI INSIEMI CANTORIANA

Lo scopo ultimo del tentativo dedekindiano di fondazione dell'aritmetica può essere considerato quello di eliminare qualsiasi riferimento psicologico (in particolare qualsiasi riferimento al tempo) dalla definizione di numero. Abbiamo visto che Dedekind inizia il suo saggio trovando nel concetto di «sistema» di «cose» qualsiasi il supporto sufficiente alla successiva introduzione dell'*Abbildung*, ossia di quell'unica facoltà della mente umana che egli richiede come necessaria e sufficiente per il suo scopo. In sostanza, la prima parte del saggio dell'88 non è che un condensato di «teoria degli insiemi» sufficiente a introdurre il concetto insiemistico fondamentale di sistema semplicemente infinito tramite il quale Dedekind caratterizza la successione dei numeri naturali.

Lo stesso Peano, nella sua assiomatizzazione, esprime l'assioma più delicato, quello di induzione, in termini di classi; e, ancora, vedremo che Frege finisce col riservare un posto fondamentale nelle sue ricerche a quelle estensioni di concetti che altro non sono che classi, insiemi, sistemi.[12] Sembra dunque non eliminabile, da questo tipo di ricerche, un riferimento più o meno essenziale, più o meno esplicito, agli insiemi e alle loro proprietà e si può dire in definitiva che la vera costante fondamentale non sia tanto quella di numero naturale quanto piuttosto quella di insieme o di classe.

In ultima istanza è come costrutto insiemistico che Dedekind introduce il sistema dei naturali, e mediante insiemi erano stati definiti – non solo da Dedekind, ma anche da Cantor e Weierstrass – gli stessi numeri reali.

Ma in che senso era possibile un discorso matematico sugli insiemi e che principi erano validi su di essi? Da questo punto di vista risulta chiaro come Georg Cantor,[13] che già più volte abbiamo indicato come il creatore di una vera e propria *teoria* degli insiemi,

[12] In questo capitolo, salvo esplicito avviso in contrario, o a meno che il contesto non indichi senza possibilità di equivoci una precisa distinzione, considereremo quindi come sinonimi termini quali «classe», «insieme», «sistema», «aggregato», «molteplicità», «estensione di un concetto» e simili.

[13] Georg Cantor nacque il 19 febbraio 1845 a Pietroburgo. Nel 1856 si sposta con la famiglia a Francoforte, in Germania; qui si rivela ben presto la sua decisa in-

vada posto di diritto nella schiera dei «fondazionisti», anche se le sue motivazioni, sulle prime, vennero più da tipici problemi matematici, al suo tempo dibattuti nell'ambito della teoria generale delle funzioni, che dalle discussioni filosofiche relative al concetto di infinito di cui ben presto fu condotto ad occuparsi.

La teoria di Cantor nasce infatti come teoria «applicata», ossia come ricerca sulle proprietà di particolari insiemi costituiti di elementi di natura ben determinata, in generale numeri o punti. E si può segnare il suo atto di concepimento, se non di nascita, già nella memoria del 1872, *Sull'estensione di una proprietà...* cui ci siamo riferiti nel paragrafo precedente. In quanto «applicata» e sempre riferita in generale a insiemi di numeri o di punti, la teoria cantoriana viene costituendosi con successivi risultati ottenuti fra il 1874 e il 1878, e presentati organicamente per la prima volta fra il 1879 e il 1884 in una serie di sei articoli dal titolo comune *Über unendliche lineare Punktmannigfaltigkeiten* (*Sulle molteplicità lineari infinite di punti*); la trasformazione sistematica in teoria astratta (ossia riferentesi a insiemi qualsiasi, o se si vuole, al *concetto* di insieme) si ha infine nei famosi *Beiträge zur Begründung der*

clinazione verso le discipline matematiche. Nel 1863 inizia a frequentare l'università di Berlino, ove si applica alla fisica, alla matematica e alla filosofia. Ebbe a maestri fra i matematici Weierstrass, Kummer e Kronecker. Sotto l'influenza di quest'ultimo si laureò nel 1867 e ottenne la libera docenza nel 1869 con lavori di tipo algebrico e di teoria dei numeri. I suoi interessi si spostarono successivamente, sotto l'influenza di Weierstrass, verso la teoria delle funzioni, concentrandosi in particolare sul problema della rappresentabilità di una funzione in serie trigonometrica. Nel 1872 viene chiamato come straordinario ad Halle; si può dire che da questo anno abbia inizio la costruzione del massimo risultato di Cantor: la teoria degli insiemi. Nel 1879 diviene ordinario, sempre ad Halle, e nel 1884 si ha la prima manifestazione di questa malattia nervosa che a più riprese dovrà colpirlo e che lo condurrà alla morte, avvenuta nella clinica psichiatrica di Halle nel 1918. Sembra probabile che al manifestarsi della malattia abbia concorso l'ostracismo scientifico e accademico decretatogli dal suo vecchio maestro Kronecker, il quale, in particolare, bloccò ogni suo tentativo di lasciare Halle per insegnare a Berlino. Abraham A. Fraenkel, uno dei maggiori studiosi contemporanei di teoria degli insiemi, in un'appendice al volume G. Cantor *Abhandlungen mathematischen und philosophischen Inhalts* (*Memorie matematiche e filosofiche*) edito nel 1932 da Ernst Zermelo, suddivide in quattro periodi la vita di Cantor (in quanto, in particolare, creatore della teoria degli insiemi): periodo della formazione (1845-71), della massima creatività (1871-84), della produttività ridotta (1884-97), della vecchiaia e del riconoscimento. Le opere di Cantor interessanti il nostro discorso verranno citate direttamente nel testo.

transfiniten Mengenlehre (*Contributi alla fondazione della teoria transfinita degli insiemi*) pubblicati in due parti nel 1895 e 1897 rispettivamente.

Il problema matematico sul quale si innestano le prime considerazioni insiemistiche di Cantor è quello della ricerca di condizioni di unicità dello sviluppo di una funzione in serie trigonometrica.[14] Abbiamo visto che nel corso della memoria del '72 – dedicata appunto a questo problema – Cantor propone fra l'altro la propria teoria degli irrazionali che ottiene modificando sensibilmente la definizione weierstrassiana. In particolare, si ricorderà, Cantor giunge a una distinzione in «ordini» dei numeri reali, in dipendenza del tipo delle rispettive successioni fondamentali cui erano associati. Orbene, proprio a questa distinzione è collegato il primo concetto «insiemistico» introdotto da Cantor (relativo sempre a insiemi di tipo particolare). Si tratta del concetto di «insieme derivato» di un insieme di punti (o, il che è lo stesso, dell'operazione di «derivazione» di un insieme di punti), che può illustrarsi come segue. Sia dato un insieme *infinito* di punti P; si dice come è noto *punto d'accumulazione* di P un punto tale che in ogni suo intorno siano contenuti infiniti punti di P. Per insieme derivato dell'insieme P Cantor intende l'insieme P' costituito da tutti i punti di accumulazione di P;[15] se P' non consiste di un numero finito (o nullo) di elementi, si potrà considerare il suo derivato P'' (derivato secondo di P) e così via; iterando questo processo per un numero *finito* n di volte, si giungerà al concetto di insieme $P^{(n)}$, derivato n-esimo di P. Per mostrare il collegamento di cui si parlava, fra i vari casi possibili (che i successivi derivati di un insieme coincidano con l'insieme stesso, che già il primo derivato contenga solo un numero finito di punti, ecc.) consideriamo il caso in cui dopo n derivazioni si giunga al derivato $P^{(n)}$

[14] Quello dello sviluppo in serie trigonometrica di funzioni (di variabile reale) anche non analitiche era un problema sollevato da Eulero (1707-83) e risolto poi da Fourier. In connessione a tale fecondissimo problema giunsero a notevoli e originali risultati matematici come Dirichlet (1805-59), Riemann, Hankel e, appunto, Cantor.

[15] Va ricordato in proposito il teorema di Bolzano-Weierstrass, secondo il quale ogni insieme infinito e limitato di punti ammette almeno un punto di accumulazione; si noti che non si richiede che un punto di accumulazione di un insieme sia necessariamente elemento dell'insieme.

dell'insieme P che contenga solo un numero finito di punti: in questo caso non essendo più possibile applicare a $P^{(n)}$ il procedimento di derivazione, si dirà che P è un insieme di *tipo* n-esimo. Orbene, un esempio di insieme di tipo n-esimo è già fornito, secondo le parole di Cantor «... da un singolo punto la cui ascissa sia un irrazionale di ordine n-esimo... Se infatti si risolve questo numero in termini di ordine $n - 1$ della successione fondamentale cui è associato e questi termini successivamente nei termini che li costituiscono di ordine $(n - 2)$-esimo e così via, si ottiene alla fine un numero infinito di numeri razionali; si pensi ora all'insieme di punti corrispondente a questi numeri, e questo insieme sarà di tipo n-esimo». È probabilmente quindi nel giusto Cavaillès quando dichiara che è proprio la teoria «nascente» degli insiemi che determina le modifiche che Cantor apporta alla definizione di Weierstrass.

Questo concetto di derivazione di un insieme di punti rappresenterà un momento centrale dello sviluppo della teoria cantoriana: supponendo di iterare il processo di derivazione al di là di ogni indice finito, Cantor è in grado di stabilire un confine al processo di passaggio al derivato partendo da un insieme chiuso e di mostrare che *ognuno* di questi insiemi si può rappresentare come riunione disgiunta di un insieme numerabile e del risultato di un'opportuna iterazione infinita del passaggio all'insieme derivato a partire dall'insieme dato.

È questo il contenuto del cosiddetto teorema di Cantor-Bendixon, che mostra l'importanza di un sistema di enti che – come i naturali – ci permettano di procedere induttivamente anche al di là del finito, nell'indiciare iterazioni di operazioni. È in questo modo che Cantor si troverà in certo senso «costretto» all'introduzione dei numeri ordinali transfiniti; e ciò avverrà, in particolare, nel 1879, nella prima delle memorie dedicate alle molteplicità infinite di punti.

Giustamente afferma dunque Zermelo che «in questo concetto di "derivazione superiore" di un insieme di punti dobbiamo ravvisare il vero e proprio "germe embrionale" e nella teoria delle serie trigonometriche il luogo di nascita della "teoria degli insiemi" cantoriana».

Siamo entrati così all'interno della problematica connessa al concetto di infinito. Dopo le significative scoperte del 1874 e del

1878,[16] Cantor comincia a intravvedere la possibilità che questi risultati possano individuare un campo autonomo di ricerca, anche se in quel periodo egli lo vedeva ancora completamente inserito nell'ambito dell'Analisi. Afferma infatti di ritenere che le generalizzazioni – riguardanti soprattutto questioni relative all'infinito – resesi a suo parere necessarie in conseguenza dei risultati ottenuti sulle varietà lineari infinite di punti «... abbiano un qualche interesse non solo per le loro applicazioni alla teoria delle funzioni, ma conducano anche a un nuovo punto di vista per quanto riguarda la conoscenza degli insiemi lineari di punti». Nel 1883, nella quinta delle memorie sopra citate *Sulle molteplicità...* (che venne anche pubblicata sotto forma di volumetto, nello stesso anno, a Lipsia, col titolo *Grundlagen einer allgemeinen Mannigfaltigkeitslehre* [*Fondamenti di una teoria generale delle molteplicità*]) egli prende netta posizione sul problema dell'infinito e lo fa ovviamente tenendo conto dei risultati ottenuti fino ad allora con la sua teoria, e in particolare riferendosi all'estensione del concetto di numero che ivi egli aveva proposto (e che noi vedremo nelle prossime pagine). Sulla stessa questione Cantor torna a più riprese, in particolare con l'articolo *Über die verschiedene Standpunkte in Bezug auf das aktuale Unendliche* (*Sui diversi punti di vista relativamente all'infinito attuale*) del 1885, e con le due *Mitteilungen zur Lehre vom Transfiniten* (*Comunicazioni sulla teoria del transfinito*) del 1887 e del 1888.

È affrontando direttamente il ruolo del concetto di infinito in matematica che Cantor è condotto non solo a prospettare una vera e propria teoria dell'infinito matematico, ma anche a sovvertire le tradizionali concezioni dei rapporti tra finito e infinito care ai matematici (e non solo ad essi) che nello stesso giro d'anni avevano trovato inaccettabile la definizione di finito a partire da quella di infinito proposta da Dedekind. A parere di Cantor l'infinito matematico può essere considerato in primo luogo nel significato di grandezza che varia, o crescendo al di là di ogni limite o diminuendo al di sotto di una quantità piccola a piacere, ma che rimane sempre *finita*; detto *infinito improprio* questa forma di infinito,

[16] Si tratta in particolare, come vedremo più avanti, dei risultati relativi alla «numerabilità» o meno di certi particolari insiemi di numeri e di punti, e alla possibilità di rappresentare biunivocamente fra loro continui geometrici con diverso numero di dimensioni.

egli osserva che nella stessa ricerca matematica recente si era in certo senso *imposta* la necessità di una diversa considerazione dell'infinito.

È importante notare che in queste prime riflessioni egli non si riferisce a propri risultati insiemistici ma, nel tentativo di vincere lo scetticismo della maggioranza dei circoli matematici (in particolare della scuola berlinese, ove primeggiava il suo antico maestro Kronecker che si era ben presto dichiarato del tutto contrario alle ricerche cantoriane), si rifà piuttosto a precise esigenze di «Analisi classica». Ad esempio, egli afferma che «... nello studio di una funzione analitica di una variabile complessa, è divenuto necessario pensare, nel piano rappresentante la variabile complessa, un unico punto posto all'infinito, ossia infinitamente distante, ma determinato, ed esaminare il comportamento della funzione in prossimità di questo punto esattamente come nelle vicinanze di qualsiasi altro punto; ne risulta che il comportamento della funzione in prossimità del punto infinitamente distante presenta le stesse "possibilità" che in ogni altro punto posto al finito, cosicché diventa pienamente legittimo in questo caso immaginare l'infinito posto in un punto completamente determinato». In casi del tipo qui esemplificato, quando cioè l'infinito si presenti come *determinato*, Cantor conviene di chiamarlo *infinito proprio* (sicché l'infinito improprio si presenta invece come «finito variabile»).

In questo concetto di infinito proprio o attuale la critica filosofica contemporanea vede, rileva Cantor, un «cattivo» infinito; questa critica non è certamente nuova, ché risale ad Aristotele, ma viene ancor oggi sostenuta con argomentazioni che possono ricondursi sostanzialmente a quelle avanzate dallo stagirita contro l'infinito attuale e in favore dell'infinito potenziale. Le due principali obiezioni aristoteliche contro l'esistenza dell'infinito «reale» vengono eliminate da Cantor riconoscendo nella prima (sostanzialmente: «Possiamo *contare* solo insiemi *finiti*») l'indebita assunzione da parte di Aristotele, che esistano solo numeri finiti: ma tale assunzione si riconduce allora a una *petitio principii*, perché è chiaro che mediante il contare possiamo riconoscere solo insiemi finiti. Ovviamente Cantor ritiene di aver proposto precisi principi per dimostrare l'esistenza, accanto a numeri finiti, anche di numeri infiniti o, come dirà, *transfiniti*. Altra ragione avanzata da Aristo-

tele contro l'infinito attuale è che se questo esistesse esso «annullerebbe» *ipso facto* il finito perché il numero finito verrebbe immediatamente «annientato» dal numero infinito. Ora le cose stanno in modo diverso, sostiene Cantor, perché è ben vero che l'aggiunzione di un numero infinito a uno finito «distrugge» in certo senso il primo, ma viceversa, l'aggiunzione di un numero finito a uno infinito può produrre una modificazione in quest'ultimo: «Questo esatto comportamento del finito e dell'infinito, totalmente trascurato da Aristotele, dovrebbe condurre a nuove suggestioni non solo nell'Analisi, ma anche nelle altre scienze, e precisamente nelle scienze naturali.»

Le condizioni e i principi posti da Cantor per l'introduzione nella matematica dei numeri transfiniti sono tali, da un lato, «... da lasciar ben poco spazio all'arbitrarietà» e d'altro lato Cantor non si stanca di ribadire che la concettualizzazione matematica ha intrinsecamente connaturato il correttivo contro assunzioni indebite o contraddittorie: essa esercita in modo quasi automatico una sorta di «rigetto» nei riguardi di quei concetti che siano sterili o contraddittori. La teoria degli insiemi si è dimostrata particolarmente feconda e va quindi assunta senza riserve nella matematica, per quanto ardita o «non standard» possa sembrare; deve essere così perché «*l'essenza della matematica sta proprio nella sua libertà*» [corsivo nostro].

Per quanto ora riguarda la concezione dell'infinito attuale esso può, secondo Cantor, presentarsi in tre forme o modalità diverse; come *assoluto*, in quanto realizzato nella sua perfezione nell'essere assoluto *extramondano*, in dio; oppure come *transfinito*, sia in quanto ricorre concretamente nel mondo dipendente delle creature, sia in quanto può essere concepito in astratto, come grandezza matematica, numero o tipo d'ordine del pensiero. L'infinito assoluto costituisce oggetto di studio della teologia speculativa, mentre l'indagine sul transfinito (in entrambe le sue forme) rientra nell'ambito della metafisica e della matematica. Nel caso comunque di accettazione dell'infinito attuale, per quanto riguarda le due forme di transfinito riconosciute da Cantor, si hanno quattro possibilità combinando in tutti i modi possibili una sua accettazione in concreto e/o una sua accettazione in astratto: Cantor si pone senza riserve nella posizione di accettazione di entrambe le modalità – astratta e concreta – dell'infinito attuale. È conscio con ciò di an-

dare contro la tradizione filosofica (e matematica), ma ritiene d'altra parte di essere il primo e il solo ad affrontare il problema con spirito nuovo e «strumentazione» scientifica adeguata.

Il primo lavoro dedicato da Cantor in modo esplicito a problemi insiemistici è una brevissima comunicazione (4 pagine) del 1874 dal titolo *Über eine Eigenschaft des Inbegriffes aller reeller algebraischen Zahlen* (*Su una proprietà dell'aggregato di tutti i numeri algebrici reali*) il cui titolo, in realtà, dice molto meno di quanto il lavoro effettivamente comprenda. In esso infatti Cantor dimostra intanto che è possibile stabilire una corrispondenza biunivoca fra l'insieme A dei numeri algebrici reali[17] (e quindi in particolare dell'insieme Q di tutti i razionali) e l'insieme N dei numeri naturali vale a dire, nella successiva terminologia cantoriana, che gli insiemi A e Q sono «numerabili». Si noti che ciò significa, sostanzialmente, che *tutti* gli elementi tanto di A quanto di Q possono porsi in una successione

$$a_1, a_2, a_3, a_4, \dots, a_n, \dots$$

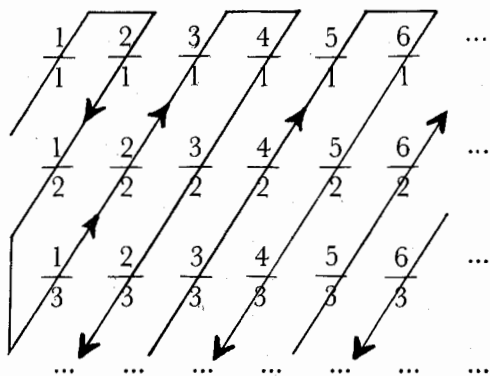
rispettivamente

$$q_1, q_2, q_3, q_4, \dots, q_n, \dots$$

dove l'indice numerico che assegna il posto di ogni elemento nella successione può essere semplicemente assunto come il numero naturale cui l'elemento in questione corrisponde. Questa dimostrazione, in effetti, non offre particolari difficoltà una volta che ci si ponga il problema; ha tuttavia notevole interesse perché l'insieme Q e l'insieme A sono entrambi *densi* (si ricordi la definizione data al paragrafo 2.3) nell'ordine «naturale» e quindi non sembrerebbe possibile poterli ordinare in modo *discreto*, ossia in modo tale che fra due elementi successivi nell'ordine dato *non* esista alcun al-

[17] Ricordiamo che una corrispondenza biunivoca fra due insiemi P e S è una trasformazione o funzione f di P in S tale che per ogni elemento $p \in P$ esista uno e un solo elemento $f(p) \in S$ e viceversa; f è *iniettiva* se per p, p' distinti, $f(p)$ e $f(p')$ sono distinti ed è invece *suriettiva* se per ogni $q \in S$ esiste un elemento $p \in P$ per cui $f(p) = q$. È immediato verificare che f è una corrispondenza biunivoca o biiettiva se è tanto iniettiva quanto suriettiva. Un numero algebrico (reale) è un numero (reale) che è radice di un'equazione algebrica a coefficienti razionali.

tro elemento dell'insieme. Tuttavia l'aver ottenuto tale dimostrazione rinforzò in un primo momento Cantor nella convinzione che in realtà vi fosse, per così dire, un solo ordine di infinito; in altri termini, un insieme o era finito o *tout court* era infinito senza che in questo secondo caso si potessero dare ulteriori specificazioni. Ma la dimostrazione, ottenuta dopo «molti sforzi e difficoltà» che l'insieme R di tutti i numeri reali *non* era associabile biunivocamente con l'insieme dei naturali, era cioè, come dirà in seguito Cantor «non numerabile» o «più che numerabile», portò alla rigorosa conclusione che si potevano considerare almeno due ordini di infinità: è proprio questo secondo importantissimo risultato che è contenuto nel lavoro qui considerato e che paradossalmente non figura nel titolo.

Trascurando per brevità la dimostrazione relativa alla numerabilità dell'insieme A, riporteremo quella relativa alla numerabilità dell'insieme Q e alla non numerabilità dell'insieme R, non seguendo però il metodo usato da Cantor in questo lavoro, ma facendo uso di metodi impiegati in un momento più maturo dell'elaborazione della teoria. Per quanto riguarda l'insieme Q si pensino i suoi elementi disposti nella matrice doppiamente infinita

di cui è evidente la legge di composizione. Si ordinino quindi i razionali della tabella seguendo la «diagonale» in essa segnata, avendo cura di eliminare dall'elencazione ogni razionale di cui sia già stato scritto un equivalente (ad esempio non scriveremo 2/2,

3/3, 4/4, ecc. una volta scritto 1/1 e così via). Si otterrà allora la successione

$$\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{1}{3}, \frac{3}{1}, \frac{4}{1}, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, \dots$$

Si noti che la tabella contiene certamente *tutti* i razionali positivi (precisamente, il numero m/n comparirà nella n-esima riga, m-esima colonna) e si noti che non è difficile estendere l'elencazione a *tutti* i numeri razionali, positivi e negativi.[18] È anche facile descrivere analiticamente l'ordine che abbiamo scelto per la dimostrazione (limitandoci per comodità ai soli numeri razionali positivi): si definisca infatti «altezza» di un numero razionale semplicemente la somma del suo numeratore e del suo denominatore. Gli elementi dell'insieme Q risultano allora ordinati «per altezza» (ossia fra due numeri diversi viene prima quello che ha altezza minore) e, a parità di altezza, per altezze pari i numeri si susseguono secondo il normale ordine di grandezza fra numeratori; per altezze dispari secondo il normale ordine di grandezza fra denominatori. La richiesta corrispondenza con l'insieme N si ottiene ora nel modo più semplice (ci limitiamo ancora ai razionali positivi) ponendo

$$\begin{array}{cccccccccc} \frac{1}{1}, & \frac{2}{1}, & \frac{1}{2}, & \frac{1}{3}, & \frac{3}{1}, & \frac{4}{1}, & \frac{3}{2}, & \frac{2}{3}, & \frac{1}{4}, & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \dots \end{array}$$

Il procedimento di dimostrazione (o più precisamente, di ordinamento) sopra esposto viene detto «procedimento diagonale di Cauchy» ed è di larghissimo uso. Ma veniamo ora alla dimostrazione della non numerabilità dell'insieme R di tutti i numeri reali

[18] Per questo è sufficiente premettere il numero zero, ad esempio nella forma 0/1, e far seguire a ogni numero della successione che così viene ottenuta il corrispondente negativo. Si avrà allora

$$\frac{0}{1}, \frac{1}{1}, -\frac{1}{1}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \dots$$

o se si preferisce, per la discussione svolta nel paragrafo precedente, del continuo lineare (anche questa volta ci serviremo di un metodo messo a punto da Cantor in una fase successiva dell'elaborazione della teoria).[19] È chiaro che basterà dimostrare che non è numerabile un sottoinsieme proprio di R, ché *a fortiori* anche R risulterà in questo caso non numerabile. Consideriamo allora il sottoinsieme R_1 dell'insieme R costituito da tutti i numeri reali compresi fra 0 e 1. Ricordiamo intanto che ogni numero reale può essere scritto in forma decimale e che che quindi l'insieme R_1 sarà composto da numeri della forma $\alpha_i = 0, a_{i1}\, a_{i2}\, a_{i3}\, a_{i4}...$ (il motivo dei doppi indici sarà subito chiaro). Tenendo inoltre conto che un numero decimale «limitato» (razionale) può essere scritto in due forme equivalenti (ad esempio 0,50000000... = 0,49999999...) conveniamo di scegliere in ogni caso una (qualsiasi) delle due (ad esempio, quella di periodo 9). Ciò premesso, la dimostrazione prosegue per assurdo. Supponiamo che il nostro insieme sia numerabile; ciò vuol dire, come sappiamo, che è possibile ordinare i suoi elementi in una successione

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \ldots$$

Supponiamo i numeri di questa successione disposti come nella tabella seguente

$$\alpha_1 = 0,a_{11}a_{12}a_{13}a_{14}a_{15}a_{16}\ldots$$

$$\alpha_2 = 0,a_{21}a_{22}a_{23}a_{24}a_{25}a_{26}\ldots$$

$$\alpha_3 = 0,a_{31}a_{32}a_{33}a_{34}a_{35}a_{36}\ldots$$

....

$$\alpha_n = 0,a_{n1}a_{n2}a_{n3}a_{n4}\ldots a_{nn}\ldots$$

....

[19] Precisamente nella nota *Über eine elementare Frage der Mannigfaltigkeitslehre* (*Su una questione elementare della teoria delle molteplicità*) del 1890.

Costruiamo ora un numero reale β compreso fra 0 e 1 e tale che, comunque la successione sia costruita, β non figuri certamente in essa. Definiamo infatti il seguente numero reale

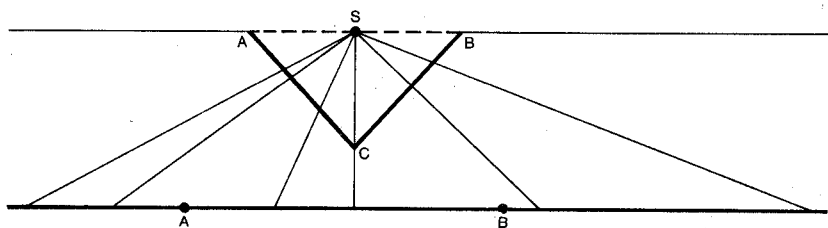
$$\beta = 0,b_1 b_2 b_3 b_4 \ldots \text{ dove } b_i = \begin{cases} 2 \text{ se } a_{ii} \neq 2 \\ 1 \text{ se } a_{ii} = 2. \end{cases}$$

È chiaro che tale numero *non* compare nella successione precedente perché differisce da α_1 almeno per la prima cifra decimale, da α_2 almeno per la seconda, in generale da α_n almeno per l'n-esima. Osservando la «freccia» tracciata nella tabella e dopo quanto detto è immediata la giustificazione del nome di «procedimento diagonale di Cantor» dato a questo metodo di dimostrazione.[20]

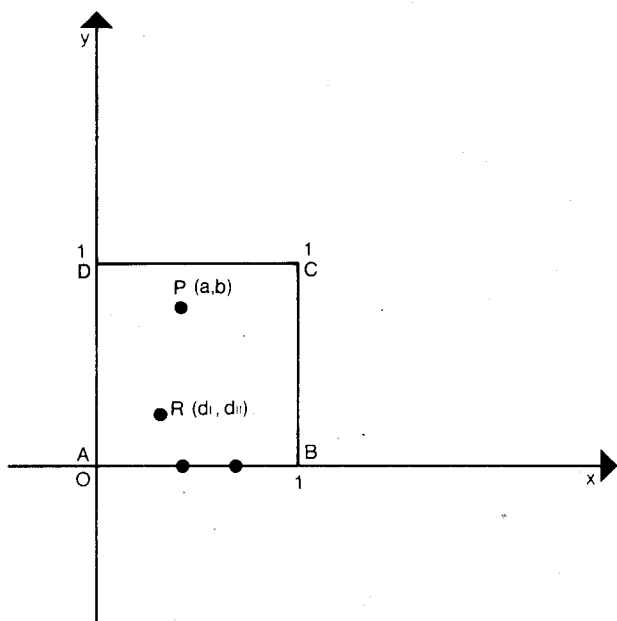
Abbiamo quindi visto che non è possibile coordinare biunivocamente già il sottoinsieme proprio R_1, dell'insieme R con l'insieme N; l'insieme R risulta pertanto «non numerabile», appartiene per così dire a un tipo di infinità diversa da quella di N, o di Q, o di A. Tenendo anzi conto del risultato sull'insieme A dei numeri algebrici, risulta subito che i numeri trascendenti costituiscono a loro volta un insieme più che numerabile.[21] Infine, dal punto di vista della caratterizzazione del continuo si è con ciò ottenuta una sua distinzione di tipo *cardinale*, rispetto ad esempio all'insieme Q dei razionali, o all'insieme N dei naturali: in generale rispetto a ogni insieme «numerabile».

[20] In effetti nella dimostrazione precedente abbiamo trascurato di fare una serie di precisazioni che, se pur a rigore necessarie, non ci sembra tocchino la sostanza dell'argomento, che viceversa verrebbe di molto appesantito con la loro introduzione. Il lettore matematico potrà ovviare da solo a queste e altre imperfezioni che lasciamo (o avremo occasione di lasciare) in vista di un'esposizione più scorrevole.

[21] Un numero reale è trascendente quando *non* è radice di alcuna equazione algebrica a coefficienti interi, ossia quando non è algebrico. Joseph Liouville aveva dimostrato nel 1844 che in ogni intervallo prefissato (α, β) della retta reale esistono infiniti numeri trascendenti; nel 1873 Charles Hermite (1822-1901) aveva dimostrato la trascendenza di e (base dei logaritmi naturali) e nel 1882 Ferdinand Lindemann dimostrerà la trascendenza di π. La dimostrazione di Cantor si pone quindi come un'affermazione di esistenza di portata più ampia di quella di Liouville, anche se non fornisce immediatamente, come quest'ultima, un metodo per *costruire* un numero trascendente, ed ha quindi, dal punto di vista della teoria dei numeri, un grado di informatività molto minore.

Altra proposizione sorprendente che si ricava immediatamente già con i metodi dimostrativi elementari sopra esposti, che si basano essenzialmente sulla possibilità o meno di stabilire corrispondenze biunivoche, è che ogni porzione di un continuo lineare può essere associata biunivocamente a tutto il continuo, o in altri termini che può stabilirsi una corrispondenza biunivoca fra *qualunque* segmento e tutta la retta reale. Allo scopo basta considerare la seguente figura, dove il segmento AB è stato «spezzato» in due parti uguali, ottenendo un triangolo isoscele, e quindi «proiettato» sulla retta r dal centro della sua base; risulta evidente che a ogni punto della retta corrisponde un punto, ed uno solo, della spezzata ACB, ossia del segmento AB.

Ma la conclusione ben più sconcertante per i matematici dell'epoca (e in effetti per lo stesso Cantor, che a proposito della proposizione che ora dimostreremo in un caso particolare ebbe a scrivere a Dedekind: «Lo vedo ma non ci credo») fu la dimostrazione che un continuo con numero di dimensioni n qualunque (finito o numerabile) può essere associato biunivocamente a un continuo lineare.

Se si ricordano i lavori di Riemann e Helmholtz in particolare, nei quali veniva assunto come caratteristica di un continuo a n dimensioni il fatto che ogni suo punto restava determinato da n variabili indipendenti, non sarà peraltro difficile capire con quanti dubbi e riserve i matematici accogliessero la dimostrazione cantoriana, che in effetti dovette attendere a lungo prima di essere pubblicata nel 1878 sul «Journal de Crelle» col titolo *Ein Beitrag zur Mannigfaltigkeitslehre* (*Un contributo alla teoria delle molteplicità*). Noi daremo qui la ingegnosa dimostrazione che risale a Cantor, limitandoci al caso di un continuo a due dimensioni (un «quadrato»).

Si consideri la figura seguente:

Si tratta di vedere che esiste una corrispondenza, fra tutti i punti del segmento AB e tutti i punti del quadrato $ABCD$ su di esso costruito. Allo scopo, assumiamo per semplicità che il lato del quadrato abbia lunghezza 1, sicché le ascisse e le ordinate di ogni punto $P(a, b)$ del quadrato (compreso il contorno) saranno del tipo $a = 0{,}a_1a_2a_3a_4...$, e $b = 0{,}\ b_1b_2b_3b_4...$ Anche qui si può assumere per convenzione una particolare notazione per i numeri nel caso di notazioni non univoche. Orbene, si prenda adesso un punto qualsiasi $P(a, b)$ del quadrato e con i due numeri reali

$$a = 0{,}a_1a_2a_3a_4a_5a_6\ ...$$

$$b = 0{,}b_1b_2b_3b_4b_5b_6\ ...$$

(coordinate di P) si costruisca un nuovo numero reale c ponendo semplicemente

$$c = 0{,}a_1b_1a_2b_2a_3b_3a_4b_4\ ...$$

la cui legge di formazione non ha bisogno di commenti. È chiaro inoltre che tale numero, essendo compreso fra 0 e 1 rappresenterà un punto E del segmento AB. Viceversa si prenda un punto F di tale segmento; a esso corrisponderà una certa ascissa reale

$$d = 0{,}d_1d_2d_3d_4d_5d_6 \dots$$

Dal numero d costruiamo altri due numeri

$$d_{\mathrm{I}} = 0{,}d_1d_3d_5d_7d_9d_{11} \dots$$

$$d_{\mathrm{II}} = 0{,}d_2d_4d_6d_8d_{10}d_{12} \dots$$

la cui legge di formazione è ancora evidente. Altrettanto evidente è che si è così individuato un punto R del quadrato, precisamente quello di coordinate $R(d_{\mathrm{I}},d_{\mathrm{II}})$.[22]

La dimostrazione di Cantor significa sostanzialmente che la «variazione» di un dato insieme di n variabili indipendenti può essere rappresentata con la variazione di una sola variabile, vale a dire che da questo punto di vista non vi è distinzione di principio fra una «varietà» continua n-dimensionale e una «varietà» continua unidimensionale. In altri termini, questo risultato poneva una prima grossa ipoteca alla ulteriore sostenibilità del concetto di dimensione come fino ad allora caratterizzato. Un altro risultato, che metterà definitivamente in crisi tale concetto, sarà ottenuto da Peano nel 1890 con la presentazione delle equazioni di una curva che «riempie» tutto un quadrato. Si osservi che in certo senso la scoperta di Peano assume qui l'aspetto di un «esperimento» che conferma un risultato (quello di Cantor) che *doveva* essere accettato in forza della inoppugnabile dimostrazione che ne era stata data, ma che per così dire non concedeva appigli intuitivi, «sperimentali». Sarà Dedekind che, prontamente, mostrerà in una lettera a Cantor come il concetto di dimensione possa essere salvato una volta che ci si restringa alle applicazioni *continue* e sia

[22] Anche qui, oltre ad aver trascurato certe puntualizzazioni a rigore necessarie, non affrontiamo il problema della natura e delle proprietà della rappresentazione, in particolare le relazioni fra biunivocità e bicontinuità della rappresentazione stessa. Infatti, come si accenna più avanti nel testo, tale rappresentazione pur essendo biunivoca non è bicontinua.

quindi una nozione topologica e non soltanto cardinale. Paradossalmente la costruzione di una teoria della dimensione lungo queste linee sarà opera, a partire dal 1910, di L.E.J. Brouwer, l'acerrimo nemico della teoria cantoriana.

Con l'articolo del '78 si può dire abbia termine la fase iniziale, preparatoria della ricerca cantoriana. Già in questo articolo in effetti viene introdotto quel concetto di «potenza» con una implicita estensione del concetto di numero al transfinito che costituirà uno dei punti centrali dell'elaborazione sistematica della teoria. Questa in effetti – come già si accennava – riceve una prima presentazione organica nei sei articoli del 1879-84 che, a detta di Zermelo, contengono la «quintessenza» di tutta l'opera creativa di Cantor: vi viene esposta infatti, anche se ancora in modo quasi completamente «applicato», la teoria dei numeri cardinali e ordinali transfiniti.

La concezione «astratta» della sistemazione finale della teoria cantoriana è subito evidente nei *Contributi* del 1895; non ci si riferisce più, qui, a insiemi particolari, finiti o infiniti, lineari o no, di punti o di numeri, ma l'articolo inizia con la famosa definizione di insieme *tout court*: «Con "insieme"», afferma Cantor, «intendiamo ogni riunione M in un tutto di oggetti m (che vengono detti "elementi" di M) della nostra intuizione o del nostro pensiero». A ogni insieme M spetta una ben determinata «potenza» o «numero cardinale» che è «... quel concetto generale... che si ottiene da M quando si astragga dalla natura particolare dei suoi elementi e dall'ordine col quale essi sono dati». Per tener conto anche simbolicamente della doppia astrazione dalla natura e dall'ordine degli elementi, Cantor, per indicare il cardinale di un insieme, pone due sbarrette sul simbolo dell'insieme stesso: così il cardinale di M, ad esempio, sarà indicato con $\overline{\overline{M}}$. Se chiamiamo «equivalenti» due insiemi M e N quando fra loro possa stabilirsi una corrispondenza (o applicazione) biunivoca, risulta che l'equivalenza è condizione necessaria e sufficiente per l'equipotenza, ossia $\overline{\overline{M}} = \overline{\overline{N}}$ allora e solo allora che M è equivalente a N.

I numeri cardinali così definiti comprendono evidentemente, come facilmente si verifica, gli usuali numeri naturali che risultano essere quei particolari cardinali associati ad insiemi *finiti*: si noti tuttavia che mai Cantor, nel corso della esposizione della sua teoria, dà una definizione di insieme finito o di insieme infinito (si

ricordi la critica di Dedekind). Fra cardinali così introdotti è possibile definire del tutto in generale le relazioni di = (uguale), ≤ (minore o uguale), ≥ (maggiore o uguale), dove $\overline{\overline{M}} \leq \overline{\overline{N}}$ se e solo se esiste una iniezione, o funzione iniettiva, $f: M \rightarrow N$; diciamo che $\overline{\overline{M}} < \overline{\overline{N}}$ se $\overline{\overline{M}} \leq \overline{\overline{N}}$ e $\overline{\overline{M}} \neq \overline{\overline{N}}$. Si può sviluppare tutta una aritmetica cardinale sulla base di opportune e naturali definizioni insiemistiche per le operazioni di addizione, moltiplicazione, esponenziazione, ecc., e in modo tale che, nel caso ci si limiti a considerare solo cardinali finiti (ossia: numeri naturali) queste relazioni e operazioni coincidano con le omonime relazioni e operazioni aritmetiche ordinarie. Le definizioni di somma, prodotto ed esponenziazione coinvolgono – come detto – operazioni su insiemi. Così, poniamo che

$$\overline{\overline{M}} + \overline{\overline{N}} = \overline{\overline{M \uplus N}}$$

ossia il cardinale della somma disgiunta di M e N, vale a dire dell'insieme ottenuto sostituendo ad esempio a M un insieme M' equivalente a M ma disgiunto da N (cioè senza elementi in comune con esso) e riunendo con N. Il prodotto $\overline{\overline{N}} \times \overline{\overline{M}}$ sarà definito invece facendo riferimento al prodotto cartesiano $M \times N = \{\langle a,b \rangle \mid a \in M \text{ e } b \in N\}$ e si avrà

$$\overline{\overline{M}} \times \overline{\overline{N}} = \overline{\overline{M \times N}}.$$

Infine $\overline{\overline{M}}^{\overline{\overline{N}}}$ sarà il cardinale di M^N definito come l'insieme di tutte le funzioni $f: N \rightarrow M$ che hanno N come dominio e prendono valori in M. È interessante notare come anche sulla base delle più generali definizioni relative a cardinali qualunque (ossia finiti o infiniti) queste operazioni godano delle usuali proprietà delle ordinarie operazioni aritmetiche. Si ha ad esempio, se indichiamo come lettere gotiche minuscole cardinali qualsiasi (cioè finiti o tranfiniti)

$$\mathfrak{a}+\mathfrak{b} = \mathfrak{b}+\mathfrak{a} \qquad \mathfrak{a}\cdot\mathfrak{b} = \mathfrak{b}\cdot\mathfrak{a}$$

$$\mathfrak{a}+(\mathfrak{b}+\mathfrak{c}) = (\mathfrak{a}+\mathfrak{b})+\mathfrak{c} \qquad \mathfrak{a}\cdot(\mathfrak{b}\cdot\mathfrak{c}) = (\mathfrak{a}\cdot\mathfrak{b})\cdot\mathfrak{c}$$

$$\mathfrak{a}\cdot(\mathfrak{b}+\mathfrak{c}) = \mathfrak{a}\cdot\mathfrak{b} + \mathfrak{a}\cdot\mathfrak{c}$$

ossia l'addizione e la moltiplicazione fra cardinali qualsiasi sono commutative e associative e inoltre la moltiplicazione è distributiva rispetto all'addizione. Anche l'esponenziazione come sopra definita relativamente a cardinali qualsiasi gode delle usuali proprietà rispetto alle altre operazioni, si ha cioè

$$\mathfrak{a}^{\mathfrak{b}} \cdot \mathfrak{a}^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}+\mathfrak{c}}$$
$$\mathfrak{a}^{\mathfrak{c}} \cdot \mathfrak{b}^{\mathfrak{c}} = (\mathfrak{a} \cdot \mathfrak{b})^{\mathfrak{c}}$$
$$(\mathfrak{a}^{\mathfrak{b}})^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b} \cdot \mathfrak{c}}.$$

Un solo comportamento «anomalo» si ha in questo grandioso tentativo di estensione e trasporto delle proprietà dell'aritmetica (cardinale) finita all'ambito generale delle potenze. Già non è immediato – ma è il contenuto di un teorema fondamentale noto come il teorema di Cantor-Bernstein – che se $\overline{\overline{M}} \leq \overline{\overline{N}}$ e $\overline{\overline{M}} \geq \overline{\overline{N}}$, allora $\overline{\overline{M}} = \overline{\overline{N}}$, cioè se esistono una iniezione di M in N e una iniezione di N in M allora esiste una biiezione tra M ed N così che $\leq$ è una relazione d'ordine fra cardinali. La vera difficoltà sorge quando ci si chiede se vale ancora, per cardinali $\mathfrak{a}$ e $\mathfrak{b}$, una qualunque delle tre relazioni $\mathfrak{a} = \mathfrak{b}$, $\mathfrak{a} > \mathfrak{b}$, $\mathfrak{a} < \mathfrak{b}$; *se* una è verificata non possono verificarsi le altre due; ma non è possibile decidere – diversamente che nel caso dei soli cardinali finiti – *se una di esse si verifichi necessariamente.* In altri termini, ci si chiede se due cardinali qualsiasi siano sempre confrontabili rispetto alla «grandezza» o, brevemente, se per cardinali qualsiasi valga la *tricotomia* e l'ordine sia *lineare.* Il problema della *confrontabilità* dei cardinali sarà, come vedremo, un'importante questione che malgrado tutti i suoi sforzi Cantor sarà costretto a lasciare aperta.

Sviluppata brevemente la teoria dei cardinali finiti, Cantor introduce il primo (che risulterà essere il minimo) cardinale transfinito, che indica con $\aleph_0$, ponendolo uguale per definizione al numero cardinale spettante all'insieme N dei naturali (cardinali finiti) cioè

$$\aleph_0 = \overline{\overline{\mathrm{N}}}.$$

Già l'introduzione di $\aleph_0$ comporta alcuni risultati «non usuali» nell'aritmetica cardinale (ferma restando ovviamente la validità delle leggi generali sopra ricordate). Ad esempio si ha

$$\aleph_0 + n = \aleph_0 + \aleph_0 = \aleph_0 \qquad (n \text{ finito})$$
$$\aleph_0 \cdot n = \aleph_0 \cdot \aleph_0 = \aleph_0 \qquad (n \text{ finito})$$
$$\aleph_0^n = \aleph_0 \qquad (n \text{ finito}).$$

Come esempio della fecondità dei concetti introdotti, Cantor mostra come egli possa derivare «in modo puramente algebrico... e in poche righe» l'intero contenuto della memoria del 1878 che tanta resistenza aveva incontrato fra i matematici. Indichiamo infatti con $\mathfrak{c}$ il numero cardinale spettante al continuo lineare (o, il che è lo stesso, la potenza dell'insieme dei numeri reali compresi fra 0 e 1, estremi inclusi). Sulla base della definizione di esponenziazione si ottiene facilmente l'uguaglianza

$$\mathfrak{c} = 2^{\aleph_0}$$

dove 2 è il cardinale dell'insieme $\{0,1\}$; per le proprietà dell'esponenziazione si ha subito:

$$\mathfrak{c}^n = \mathfrak{c}$$

per n finito, ma anche

$$\mathfrak{c}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0} = \mathfrak{c}$$

ossia: la potenza di un continuo n-dimensionale, per n finito o numerabile, è uguale alla potenza del continuo lineare. Come ultima osservazione, ricordiamo che Cantor, generalizzando quanto provato sui rapporti fra naturali e reali, aveva dimostrato col metodo diagonale già nel 1890 un teorema (oggi noto semplicemente come *teorema di Cantor*) il quale in termini cardinali afferma che

$$2^{\mathfrak{m}} > \mathfrak{m}$$

per $\mathfrak{m}$ cardinale *qualsiasi*. Sulla base di questo fondamentale teorema,[23] a partire da $\aleph_0$ è possibile allora ottenere una successione infinita di cardinali della forma

$$2^{\aleph_0} \quad 2^{2^{\aleph_0}} \quad 2^{2^{2^{\aleph_0}}} \qquad \dots \text{ecc.}$$

[23] Dato un insieme M, di cardinale $\overline{\overline{M}}$, e indicato con $\mathscr{P}(M)$ l'insieme di tutti i sottoinsiemi di M, si dimostra facilmente che

$$\overline{\overline{\mathscr{P}(M)}} = 2^{\overline{\overline{M}}}.$$

ognuno dei quali è maggiore del precedente. Su questo modo di generare cardinali torneremo nel seguito.

I cardinali cantoriani generalizzano i cardinali finiti, ma anche gli ordinali possono avere una controparte infinita. Cantor basa la caratterizzazione *ordinale* degli insiemi sul concetto di *tipo d'ordine.* Tipo d'ordine di un insieme M è «... il concetto generale che si ottiene da M quando si astrae soltanto dalla natura degli elementi m, ma non dall'ordine con cui sono dati». Questa «definizione» rende ragione della notazione $\overline{M}$ scelta da Cantor per indicare il tipo d'ordine dell'insieme M. Dati due insieme ordinati M e N (ossia su ognuno dei quali sia definita una relazione d'ordine lineare,[24] che indicheremo, all'occorrenza, con $<_M$ e $<_N$ rispettivamente) a ognuno di essi spetterà un determinato tipo d'ordine, $\overline{M}$, rispettivamente $\overline{N}$, e il criterio di uguaglianza fra tipi d'ordine, $\overline{M} = \overline{N}$, diventa ora la *similitudine,* ove si intenda che gli insiemi M e N sono simili se fra di essi può stabilirsi una corrispondenza f tale che

Basterà osservare la corrispondenza biunivoca fra i sottoinsiemi $Y \subseteq M$ e le *funzioni caratteristiche* $f_Y : M \to \{0,1\}$ definite ponendo

$$f_Y(x) = \begin{cases} 0 \text{ se } x \notin Y \\ 1 \text{ se } x \in Y. \end{cases}$$

In termini insiemistici allora il teorema di Cantor afferma che la cardinalità dell'insieme $\mathscr{P}(M)$ di tutti i sottoinsiemi di un insieme M dato è maggiore della cardinalità dell'insieme M, e ciò vale qualunque sia la cardinalità, finita o infinita, di M.

[24] Ricordiamo che una relazione R si dice un *ordine* (o, come spesso viene chiamato, un *ordine parziale*) se è:
1) Riflessiva, cioè $\forall x(xRx)$
2) Transitiva, cioè $\forall x \forall y \forall z(xRy \wedge yRz \to xRz)$
3) Antisimmetrica, cioè $\forall x \forall y(xRy \wedge yRx \to x = y)$.
Si parla di ordine *stretto* nel caso di una R che goda di 2) ma sia
4) Irriflessiva, cioè $\forall x(\neg\, xRx)$.
Spesso si indica in generale con $<$ un ordine stretto e con $\leq$ un ordine. C'è una corrispondenza biunivoca fra i due concetti, in quanto avremo in generale che:

$$x \leq y \text{ se } x < y \vee x = y,$$
$$x < y \text{ se } x \leq y \wedge x \neq y.$$

Un ordine si dice *totale* o *lineare* se gode della proprietà
5) $$\forall x \forall y(x \leq y \vee y \leq x).$$
La 5) si traduce immediatamente nella proprietà di *tricotomia* o *connessione* per ordini stretti
6) $$\forall x \forall y(x \leq y \vee x = \vee y < x).$$

a) f è biunivoca;[25]
b) f «conserva» l'ordine, ossia: se a, b sono due qualsiasi elementi di M e $f(a)$, $f(b)$ indicano i corrispondenti elementi di N, se si ha $a <_M b$, si ha anche $f(a) <_N f(b)$ e viceversa.

Sulla base del concetto di tipo d'ordine si può così stabilire – in parallelo col caso cardinale – tutta un'aritmetica ordinale che ancora, nel caso ci si limiti a considerare tipi d'ordine di insiemi finiti, coincide con l'ordinaria aritmetica ordinale dei numeri naturali, e quindi con la corrispondente aritmetica cardinale; nel caso però di tipi d'ordine *qualsiasi* (ossia in generale di insiemi anche infiniti) si hanno ora notevoli divergenze col caso ordinario: così, ad esempio, le operazioni di addizione e moltiplicazione, opportunamente definite in modo naturale, non godono più, in generale, della proprietà commutativa.

Per avere un concetto di numero ordinale adeguato agli insiemi infiniti e che rifletta l'idea intuitiva di «indicatore di posto» in una lista, non basta però limitarsi ai generici insiemi ordinati. Assume così un'importanza fondamentale nella teoria di Cantor una sottoclasse particolare dei tipi d'ordine: si tratta dei tipi d'ordine di insiemi *bene ordinati* (tali cioè che sono ordinati e inoltre ogni loro sottoinsieme non vuoto ha un primo elemento rispetto alla relazione ordinatrice) che vengono chiamati da Cantor *numeri ordinali* o semplicemente *numeri*. Essi costituiscono l'estensione «canonica» del concetto di numero ordinale dal finito al transfinito e rappresentano per Cantor il «materiale naturale per una adeguata definizione delle cardinalità o potenze superiori». I numeri ordinali inoltre godono della tricotomia, ossia si può *dimostrare* che, per α, β ordinali qualunque, delle tre relazioni $\alpha = \beta$, $\alpha > \beta$, $\alpha < \beta$, non solo ognuna di esse esclude le altre due ma anche una delle tre ha necessariamente luogo. In altri termini, per i numeri ordinali non esiste un «problema di confrontabilità» analogo a quello visto per i cardinali.

In modo analogo alla teoria dei cardinali, anche sugli ordinali è possibile definire operazioni quali somma, prodotto, esponenzia-

[25] Quindi, in particolare, se due insiemi hanno lo stesso tipo d'ordine essi hanno anche lo stesso numero cardinale. Il viceversa si verifica sì nel caso di insiemi finiti, ma non vale in generale: dato un insieme il cui cardinale sia un qualunque numero transfinito, esistono infiniti tipi d'ordine che gli corrispondono. Si confronti anche la nota 27 a pag. 324.

zione, ecc., che si definiscono in termini di «operazioni» fra insiemi ordinati. Ad esempio, se consideriamo il tipo d'ordine ω dei naturali, possiamo definire il suo successore $\omega + 1$ come il tipo d'ordine dell'insieme bene ordinato N* che si ottiene aggiungendo dopo *tutti* i naturali un nuovo elemento. L'aritmetica che così risulta è molto più sottile di quella cardinale e consente ad esempio la formulazione di un concetto generale di sviluppo che permette di dare una forma normale di base ω per tutti i numeri ordinali, mostrando come l'aritmetica ordinale offra molte analogie con l'aritmetica finita (si pensi, a proposito del risultato di sopra, allo sviluppo di ogni numero naturale in base 10, o 2, ecc.) ma anche significative differenze. Ad esempio non vale la proprietà commutativa di $+$ e $\times$; così si ha $1 + \omega = \omega \neq \omega + 1$ e $2 \cdot \omega = \omega \neq \omega \cdot 2 = \omega + \omega$. Su questa base Cantor rispose alla seconda obiezione di Aristotele che abbiamo visto a pagina 307.

Centrale nell'aritmetica ordinale, è il ruolo del principio di *induzione transfinita* che generalizza l'induzione sui numeri naturali e traduce a livello inferenziale il fatto che – come l'insieme dei naturali – anche la classe di tutti gli ordinali, come ogni suo segmento, è bene ordinata, vale a dire, equivalentemente, ogni insieme non vuoto di ordinali ha un primo elemento, e ogni insieme di ordinali ha un sup. Nel caso dei naturali sappiamo che l'induzione si può formulare dicendo che se una proprietà P vale per 0 e se preso un qualunque x se P vale per x vale anche per il suo successore $x + 1$, allora P vale per ogni naturale. Questo dipende dal fatto che ogni naturale si ottiene da 0 applicando un numero finito di volte l'operazione di successore. Nel caso degli ordinali transfiniti, si può dare una formulazione analoga del principio di induzione, ma non basta considerare la sola operazione di successore. È chiaro infatti che già il primo ordinale transfinito, ω, non è il successore di alcun ordinale finito, in quanto appunto il successore di ogni numero naturale è ancora un numero naturale. Chiamiamo *limite* gli ordinali $\neq 0$ che, come ω, non sono successori di alcun ordinale. In generale ogni ordinale limite si può ottenere come risultato di un'operazione infinita di *passaggio al limite*, considerando gli ordinali strettamente minori e definendo, se $\{\alpha_\beta\}$ per $\alpha < \lambda$ è una successione crescente di ordinali, che lim $\{\alpha_\beta\}$ è il più piccolo ordinale maggiore di ogni α_β per $\beta < \lambda$. Che questo ordinale esista scende dal fatto che la classe degli ordinali è bene ordinata e possiamo concludere in generale che

ogni ordinale $\neq 0$ o è successore di un ordinale precedente o è limite di una successione crescente di ordinali minori. Possiamo così formulare il principio di induzione transfinita dicendo che se P è goduta da 0 ed è conservata dal passaggio al successore e ai limiti allora P è goduta da tutti gli ordinali. Avremo occasione più avanti di ritornare sugli ordinali accennando – tra l'altro – al loro ruolo nella teoria della dimostrazione in particolare dopo Gentzen.

Per Cantor il ruolo precipuo degli ordinali è quello di organizzare lo studio dei cardinali transfiniti. Vediamo quindi come Cantor definisce le «cardinalità superiori» tramite i numeri ordinali. Pensiamo di riunire in una prima *classe numerica* tutti i numeri ordinali (di insiemi bene ordinati) finiti: questa classe (che come sappiamo coincide con l'insieme N dei numeri naturali), pensati i suoi elementi ordinati secondo l'ordine di grandezza naturale (relazione di «maggiore»), è caratterizzata dall'avere un primo elemento, 0, ma non un ultimo elemento; dall'essere bene ordinata dalla relazione di «maggiore» e dall'avere, in quanto classe, un numero cardinale $\aleph_0$, che è maggiore di ogni elemento della classe.[26] Si riuniscano ora in una *seconda classe numerica* tutti i numeri ordinali (di insiemi bene ordinati) il cui cardinale sia uguale ad $\aleph_0$ (ossia uguale al cardinale della *classe* numerica precedente). Cantor dimostra che anche questa seconda classe ha un primo elemento (per il quale egli usa in un primo momento il simbolo ω, successivamente il simbolo ω_0, e che per definizione è l'ordinale dell'insieme dei numeri naturali ordinati secondo l'usuale ordine di grandezza, ossia $\omega_0 = \overline{N}$) ma non un ultimo elemento; è bene ordinata dalla relazione di «maggiore» ed è tale che il suo numero cardinale non solo è maggiore del cardinale di ogni suo elemento (quindi di $\aleph_0$) ma è anche il *più piccolo* cardinale transfinito che sia maggiore di $\aleph_0$; Cantor indica tale cardinale con $\aleph_1$.[27] Ora è chiaro

[26] In realtà, contrariamente a quanto avverrà per le classi numeriche successive (si veda avanti nel testo) gli elementi della prima *non* godono della proprietà di avere una stessa cardinalità, sicché in effetti ogni numero ordinale finito costituisce, singolarmente preso, una classe numerica. Si veda anche la nota successiva.

[27] Vediamo di renderci intuitivamente conto del fatto che a un insieme infinito di data cardinalità possono corrispondere più tipi d'ordine e, in particolare, più ordinali fra loro diversi. Rifacciamoci allo scopo a un esempio che già abbiamo usato quando abbiamo stabilito la numerabilità dell'insieme Q dei razionali. Sostanzialmente – come si ricorderà – la dimostrazione si basava sul fatto che si riusciva a *ordinare* diversamente, rispetto all'ordine naturale di grandezza, gli ele-

come iterare il procedimento: alla terza *classe numerica* composta da tutti gli ordinali (di insiemi bene ordinati) di cardinalità $\aleph_1$ spetterà un nuovo numero cardinale che risulterà essere maggiore di $\aleph_1$ e anzi risulterà essere il più piccolo cardinale transfinito con questa proprietà; esso verrà indicato con $\aleph_2$ e così via. In questo modo Cantor ottiene una successione transfinita di alef:

$$\aleph_0, \aleph_1, \aleph_2, \ldots, \aleph_{\omega_0}, \aleph_{\omega_0+1}, \ldots$$

che egli ritiene coincidere con quella di tutti i cardinali.

In effetti Cantor non conduce questa dimostrazione e anzi nella sua costruzione si «ferma» alla considerazione della seconda classe (malgrado abbia ovviamente gli strumenti per spingersi oltre). È probabile che ciò dipenda dal fatto che già nel 1895 Cantor si era imbattuto in una «antinomia», detta del *massimo numero ordinale* che Cesare Burali-Forti aveva poi pubblicato nel 1897, relativa alla classe bene ordinata (secondo l'ordine naturale di grandezza) Ω di tutti i numeri ordinali e che sostanzialmente si può esprimere dicendo che il numero ordinale $\overline{\Omega}$ spettante a Ω sulla base di pre-

menti di Q (vale a dire: si definiva in esso una *relazione d'ordine* diversa da quella usuale di «maggiore»). Orbene, dovrebbe anche risultar chiaro che così facendo si era cambiato il *tipo d'ordine* dell'insieme Q (contrariamente a quanto avviene nel caso di un insieme finito, ove la diversa disposizione degli elementi *non* cambia il tipo d'ordine dell'insieme): e infatti, ad esempio, nell'ordinamento naturale per grandezza Q è denso mentre il nuovo ordine lo rendeva discreto; nell'ordine naturale Q non aveva né primo né ultimo elemento, mentre col nuovo ordine rimaneva sempre senza ultimo elemento, ma aveva un primo elemento e così via. In generale quindi – parlando intuitivamente – uno stesso insieme infinito si può ordinare in infiniti modi diversi cui corrispondono infiniti tipi d'ordine diversi. Sempre intuitivamente una parte – ancora infinita – di questi ordini saranno buoni ordini e quindi, corrispondentemente, una parte di quei tipi d'ordine saranno numeri ordinali. Dovrebbe essere chiaro come si possa parlare di «cardinale» di un ordinale: basta pensare a un insieme che abbia quel cardinale e quell'ordinale. Così ad esempio si ha subito, conservando la notazione cantoriana, $\overline{\overline{\omega}}_0 = \aleph_0$. Per quanto riguarda la seconda classe cantoriana, i suoi elementi, ognuno dei quali è *numerabile*, saranno nell'ordine

$$\omega, \omega + 1, \omega + 2, \ldots \omega \cdot 2, \omega \cdot 2 + 1, \ldots$$

$$\omega \cdot 3, \omega \cdot 3 + 1, \ldots \omega^2, \omega^2 + 1, \omega^2 + 2, \ldots \omega^2 + \omega, \ldots \omega^3, \ldots \omega^4, \ldots \omega^n, \ldots \omega^\omega, \ldots$$

e così via. Come nell'aritmetica ordinale finita si conta il primo, il secondo, ecc. elemento di un insieme, ora, con la concezione dell'infinito attuale, conteremo anche l'ω-esimo, l'$(\omega + 1)$-esimo elemento e così via.

cisi risultati della teoria cantoriana, dovrebbe contemporaneamente essere e *non* essere l'ordinale di Ω (più semplicemente: risulta essere e contemporaneamente non essere il *massimo* numero ordinale).[28] Contrariamente a Burali-Forti che vide in questo fatto una difficoltà del concetto di ordinale transfinito, Cantor – che distingueva fra molteplicità coerenti e non – sfruttò *positivamente* la possibilità di bene ordinare la classe degli ordinali nel tentativo di dimostrare quel principio del buon ordinamento (ogni insieme ammette una relazione d'ordine che è un buon ordine) di cui solo nel 1904 Zermelo – come vedremo – avrebbe pubblicato una dimostrazione rigorosa e che per Cantor era essenziale.

Abbiamo parlato di distinzione fra molteplicità coerenti e non. A questa distinzione Cantor fu portato dalla scoperta di un'altra antinomia, detta del massimo numero cardinale, che aveva comunicato a Dedekind in una lettera del 1899. Pensiamo all'insieme V di *tutti* gli insiemi; si può facilmente dimostrare che ad esso spetta il *massimo* numero cardinale possibile, che conveniamo di indicare con $\aleph$. D'altra parte consideriamo l'insieme $\mathscr{P}(V)$ di *tutti* i sottoinsiemi di V; come sappiamo si ha

$$\overline{\overline{\mathscr{P}(V)}} = 2^{\overline{\overline{V}}}$$

e per il teorema di Cantor risulta

$$2^{\overline{\overline{V}}} > \overline{\overline{V}},$$

sicché essendo $\overline{\overline{V}} = \aleph$, $\aleph$ *non* è il massimo numero cardinale. L'unica soluzione possibile – per Cantor – era quella di distinguere tra totalità: quelle che possiamo considerare come *oggetti* (molteplicità coerenti) e che possiamo raccogliere in nuove collezioni, e quelle che, come V, non possiamo considerare come tali (molteplicità incoerenti). È in sostanza la distinzione fra classi e insiemi che più tardi avrebbe sviluppato von Neumann.

Come dicevamo è probabile che queste considerazioni inducessero Cantor a osservare una certa cautela nel parlare dell'«insieme» di *tutti* gli alef o di *tutti* gli ordinali. Egli non diede un peso essenziale a

[28] Si dimostra infatti: *a*) che Ω è un insieme bene ordinato; *b*) che ogni ordinale α di Ω è l'ordinale spettante all'insieme bene ordinato degli ordinali che lo precedono («segmento» di α); che un segmento di Ω non è mai simile a Ω. Da ciò si deriva immediatamente un'antinomia, appunto applicando tali risultati a $\overline{\Omega}$ e Ω.

queste antinomie che toccavano i fondamenti *logici* della teoria ma non il suo contenuto *matematico* e il suo problema centrale rimase sempre quello della confrontabilità dei cardinali, e una questione ad esso in certo senso subordinata, il cosiddetto *problema del continuo*.

Ecco come Cantor stesso si esprime nella citata lettera a Dedekind del 1899: «Come Lei sa sono giunto già molti anni fa a definire una successione bene ordinata di potenze... che chiamo "alef"... La grossa questione era se, oltre agli alef esistessero altre potenze di insiemi; già da due anni sono in possesso di una dimostrazione del fatto che non ne esistono altre, cosicché ad esempio al continuo reale aritmetico (la totalità dei numeri reali) spetta come cardinale un alef determinato». In questo passo sono chiaramente enunciati il problema generale della confrontabilità e quello particolare del continuo, sui quali è opportuno soffermarci brevemente.

Il problema della confrontabilità consiste nel chiedersi (una volta individuato, grazie alla teoria degli ordinali, un metodo «standard» per costruire potenze sempre maggiori) se ogni altro possibile metodo di generazione di cardinali conduca necessariamente a individuare sempre e solo alef. Ad esempio, noi conosciamo già un modo diverso di generazione, quello basato sull'applicazione del teorema di Cantor: orbene, ogni cardinale ottenuto tramite «esponenziazione» coincide con qualche alef? Supponiamo ora, per un momento, che l'affermazione di Cantor risponda a verità, cioè che egli sia riuscito effettivamente a dimostrare un teorema di confrontabilità, sicché ogni cardinale sarebbe un alef. A questo punto interviene l'altro problema. Abbiamo infatti visto che, in particolare, il continuo lineare geometrico ha un numero cardinale $\mathfrak{c} = 2^{\aleph_0}$; per il risultato sopra ammesso, $2^{\aleph_0}$ deve ora coincidere con qualcuno degli alef. Cantor aveva dimostrato che $\aleph_1$ è il cardinale *immediatamente* successivo a $\aleph_0$ in ordine di grandezza naturale, sicché sorge la questione di sapere se $2^{\aleph_0}$ è maggiore o uguale a $\aleph_1$: è questo appunto il *problema del continuo*. A questo problema Cantor risponde formulando la cosiddetta *ipotesi del continuo* che, sempre nel caso più semplice, si esprime affermando l'uguaglianza[29]

$$2^{\aleph_0} = \aleph_1.$$

[29] L'ipotesi generalizzata del continuo viene invece espressa dall'uguaglianza generale $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ che si afferma valere per ogni ordinale α.

Cantor tentò vanamente di dimostrare questa ipotesi, escogitando reiteratamente nuove dimostrazioni per il problema della confrontabilità, che tuttavia presentavano sempre nuovi punti di oscuro ricorso all'«intuizione» e che quindi non lo soddisfacevano. Mentre quest'ultimo veniva risolto agli inizi del XX secolo (precisamente nel 1904) da Zermelo, con l'impiego dell'assioma di scelta, la dimostrazione dell'ipotesi del continuo ha impegnato logici e matematici fino al 1963, anno in cui, come meglio vedremo più avanti, Paul J. Cohen giunse a una definita conclusione circa la sua dimostrabilità o meno. E l'esito di tali ricerche non è stato certamente quello che Cantor immaginava di poter ottenere.

Dopo quanto visto sopra sull'aspetto puramente matematico della teoria è opportuno soffermarsi ora sui problemi logici che essa sollevava. È chiaro che con essa Cantor voleva *anche* fornire una fondazione per la matematica, e si inseriva cioè nella problematica cui ci interessiamo in modo particolare in questo capitolo, ma voleva anche fornire una fondazione di tipo generale, che si potesse considerare riferita all'accezione più ampia possibile di «aritmetica» nella quale non fosse lasciato alla sola intuizione il dato ultimo della sua costruzione, quel concetto di insieme che va quindi considerato in tutta la sua oggettività matematica. Nella pratica operativa il «classico» Dedekind era indubbiamente molto più esigente, intransigente e rigoroso che non Cantor; il «romantico» Cantor, tuttavia, riusciva ad aprire nuovi orizzonti alla ricerca e filosofica e matematica, nella sua convinzione che al di là dell'esigenza immediata di rigore fosse più importante e vitale mantenere alla ricerca astratta la totale *libertà* d'espressione, fondata, oltre che sulla «sicurezza» del procedimento tecnico su una sorta di ingenua fede nelle possibilità di «autocorrezione» del pensiero matematico.

La teoria di Cantor pur se non formulata in modo sistematico dal suo creatore, si basava su tre principi che col tempo sarebbero stati isolati e formulati in modo esplicito: 1) due insiemi che hanno gli stessi elementi sono uguali (*principio di estensionalità*); 2) data una qualunque proprietà esiste sempre (ed è unico in virtù del principio precedente) l'insieme di tutti e soli gli «oggetti» che godono di quella proprietà (*principio di comprensione*); e infine 3) è sempre possibile bene ordinare un insieme di cardinalità arbitraria di insiemi non vuoti (*principio del buon ordinamento*). Di questi tre principi Can-

tor enuncia esplicitamente (anche se non consapevolmente come principio generale) solo il primo; degli altri due si serve ogniqualvolta gli occorra in modo intuitivo e inconsapevole.

Il primo principio non è particolarmente problematico; esso sta a significare sostanzialmente che un insieme è completamente determinato dai suoi elementi, indipendentemente dal fatto che possa essere caratterizzato tramite proprietà diverse e separa così nettamente la nozione *estensionale* di insieme da quella *intensionale* di proprietà.

Il terzo principio è particolarmente importante nella teoria cantoriana perché proprio ad esso, ad esempio, Cantor si riferisce inconsapevolmente nelle sue «dimostrazioni» del teorema di confrontabilità,[30] basandosi sulla distinzione fra molteplicità coerenti e incoerenti. Come si ricorderà, Cantor nella lettera sopra citata a Dedekind era giunto alla conclusione che tanto la classe di tutti gli ordinali, quanto la classe di tutti gli alef sono in effetti molteplicità *incoerenti*, ossia tali che «l'ipotesi di un "essere assieme" di *tutti* i loro elementi conduce a una contraddizione sicché è impossibile concepire la molteplicità come un'unità, come una "cosa compiuta"».[31] Di questa distinzione e della contraddizione che ne risulterebbe eliminandola, Cantor fa un uso positivo, cercando di rispondere alla domanda se esista un insieme la cui cardinalità non sia un alef che è ovviamente un altro modo di formulare il proble-

[30] Nel 1915 F. Hartogs dimostrerà che il problema della confrontabilità, il teorema del buon ordinamento e l'assioma della scelta (si veda avanti nel testo) sono proposizioni equivalenti. Zermelo aveva dimostrato nel 1904 il teorema del buon ordinamento facendo uso in modo essenziale dell'assioma di scelta (sicché il principio della scelta veniva assunto *esplicitamente* fra le proposizioni matematiche fondamentali). Quest'ultimo era stato notato «di passaggio» da Peano già nel 1890 ed enunciato come principio indipendente da Beppo Levi nel 1904. Poche proposizioni matematiche hanno provocato tante accese discussioni come questa. Avremo occasione di parlarne più avanti.

[31] In altri termini, dal momento che la teoria riguarda «insiemi» e non «molteplicità» qualunque, le antinomie non hanno più luogo. Così facendo Cantor anticipa di fatto una delle vie poi seguite per uscire dalla cosiddetta *crisi dei fondamenti della matematica* (si veda il prossimo capitolo) causata proprio dalla scoperta di antinomie nella teoria degli insiemi, in particolare dall'antinomia di Russell; in questo tipo di soluzione si dirà tecnicamente «classe» per quelle che Cantor chiamava «molteplicità incoerenti» mentre si riserva il nome di «insieme» alle cantoriane «molteplicità coerenti». Vedremo meglio la cosa in un prossimo capitolo.

ma della confrontabilità, in quanto se tutti i cardinali fossero alef sarebbero cardinali di ordinali e quindi per definizione confrontabili. Ragionando per assurdo egli «dimostra» che se *V* fosse una «molteplicità determinata alla quale *non spetta come numero cardinale un alef... V* dovrebbe essere *incoerente*»: contro l'ipotesi. È proprio per dimostrare questo fatto egli fa un uso essenziale del principio del buon ordinamento che egli evidentemente considerava come implicito nello stesso concetto di insieme.

Ma, almeno dal punto di vista strettamente fondazionale è il secondo principio, il principio di comprensione, che, pur non risultando problematico per Cantor, porterà a difficoltà radicali per la teoria. In effetti sarà proprio un'ennesima antinomia scoperta da Russell e che chiamava direttamente in causa questo principio, a costringere a rivedere tutta l'impostazione logica della teoria degli insiemi. Va detto che l'antinomia in questione, la celebre *antinomia di Russell*, riguardava – come vedremo – più il concetto di proprietà che quello di insieme. Come abbiamo visto infatti Cantor era perfettamente consapevole della necessità di distinguere fra molteplicità e in quest'ottica il paradosso di Russell, come tutti gli altri paradossi, non toccava più di tanto la pratica insiemistica. L'antinomia colpiva la teoria indirettamente in quanto mostrava la necessità, una volta che si volesse dare una *fondazione* alla teoria del transfinito, di individuare *precisamente* le proprietà che davano origine ad insiemi eliminando ogni discorso *ad hoc*, caso per caso, come tendeva a fare Cantor e la maggior parte degli insiemisti dopo di lui. In un'ottica fondazionale era necessario porsi la domanda di quali proprietà potevano occorrere nel principio di comprensione e come si poteva fare per determinarle. Mentre nel caso del principio di scelta si trattava dell'*esplicitazione* di un fondamento sul quale era inconsapevolmente basata la teoria cantoriana, per il principio di comprensione si trattava piuttosto della necessità di una sua *limitazione*, pena appunto la sostanziale indeterminatezza della teoria medesima.

In effetti, altri punti della teoria degli insiemi, così come concepita ed elaborata da Cantor, lasciavano a desiderare dal punto di vista del rigore logico. Ad esempio la stessa definizione cantoriana del concetto di insieme è evidentemente circolare; la definizione di numero cardinale o ordinale è basata su una concezione ingenua e psicologistica dell'astrazione, sicché a ben ragione Frege potrà di-

re che Cantor richiede al lettore «astrazioni impossibili». Ancora, Cantor non si preoccupa mai di definire cosa intenda per insieme finito, eppure sull'insieme di tutti i numeri *finiti* (e quindi appunto, in sostanza, sul concetto di insieme finito) fonda tutta la sua aritmetica transfinita, a partire dalla definizione che egli dà di $\aleph_0$; egli impiega acriticamente l'induzione completa, senza preoccuparsi minimamente di giustificare questo procedimento e pur essendo certamente a conoscenza delle opere, se non di Frege, almeno di Dedekind.

Malgrado questa indeterminatezza che contribuì ad acuire l'isolamento in cui sulle prime la teoria si sviluppò, non furono pochi i matematici di prim'ordine che accettarono la sfida che la teoria del transfinito presentava, ottenendo risultati estremamente interessanti. Tra gli altri possiamo ricordare Julius König (1849-1913) e Dienes König (creatore della teoria dei grafi infiniti), Félix Bernstein (1878-1956), Philip Jourdain (1879-1919), A. Schönflies che contribuì in particolare alla teoria degli insiemi di punti, ed infine Félix Hausdorff che dette per primo la definizione astratta di spazio topologico su basi insiemistiche che oggi è al centro della topologia generale e mostrò con diversi risultati l'importanza della teoria degli insiemi nello sviluppo della matematica (fondamentale fu la pubblicazione del suo volume *Grundzüge der Mengenlehre* [*Lineamenti della teoria degli insiemi*], nel 1914). Anche in Francia ed in Inghilterra (oltre a Jourdain è importante ricordare i lavori fondamentali dei coniugi W.H. e G.Ch. Young) la teoria ebbe sulle prime una certa fortuna nel tumultuoso sviluppo della teoria delle funzioni e di quella della misura per opera di René Baire (1874-1932), Henri Lebesgue (1875-1941), Henri Poincaré (1845-1912), ecc., ma presto non sarebbero mancate le resistenze, motivate in particolare da differenti concezioni sulla filosofia della matematica. Sarà in Polonia (e in misura minore in Russia), che la teoria degli insiemi diventerà, agli inizi del secolo, la base di un nuovo modo di concepire la matematica in termini astratti e non più computazionali, rimodellando Analisi, algebra, topologia, in maniera radicale.

Malgrado i limiti obiettivi che la teoria di Cantor presenta da un punto di vista del rigore logico, essa ha costituito il punto di riferimento per la formulazione astratta di gran parte della matematica contemporanea, ha avuto notevolissimi riflessi nello stesso

ambito filosofico e di fatto – come vedremo nel corso della nostra storia – il confronto positivo o negativo con la teoria degli insiemi è un elemento costante dello sviluppo della logica in questo secolo: come ebbe a dire Hilbert nel 1925 «nessuno potrà scacciarci dal paradiso che Cantor ha creato per noi».

4. LA SISTEMAZIONE DELLA LOGICA MODERNA E LA LOGICIZZAZIONE DELLA MATEMATICA: GOTTLOB FREGE

Non è possibile comprendere compiutamente il senso della svolta impressa alla logica moderna da Gottlob Frege,[32] se la si considera separata dal contesto generale nel quale fu concepita, e entro il quale solo è possibile valutarne la portata e i limiti.

Il disegno complessivo di tutta l'opera di Frege si può riassumere in quello che oggi viene detto «programma logicista» di fondazione della matematica, che a lui appunto risale e che si concretizza in due obiettivi fondamentali:

a) *definire* in termini puramente logici i concetti della matematica pura, in particolare, quelli tradizionalmente e comunemente

[32] Friedrich Gottlob Frege nacque a Wismar, nella Germania settentrionale, l'8 novembre 1848. Compì i primi studi nel locale ginnasio, quindi frequentò l'università di Jena nel 1869 e nel 1871 quella di Gottinga, ove ebbe a maestro fra gli altri Hermann Lotze, dal quale con tutta probabilità trasse l'impostazione platonico-realista in filosofia. Laureatosi nel 1873 a Gottinga, conseguì la libera docenza nel 1874. Nel 1879 viene nominato professore straordinario all'università di Jena presso la quale compie tutta la sua carriera accademica che si conclude nel 1895 con la nomina a professore ordinario onorario. Nel 1917 chiede alle autorità accademiche di essere esonerato dall'insegnamento e nel 1918 viene collocato a riposo. Morì il 25 luglio 1925. Il carattere di Frege, già chiuso e polemico per natura, si inasprì probabilmente ancor più per la totale indifferenza con la quale gli ambienti scientifici tedeschi accolsero le sue varie opere dedicate, dal 1879 in poi, all'esecuzione del proprio «programma logicista» di fondazione della matematica. Oltre alle opere inerenti questo programma, che verranno citate direttamente nel testo, ricordiamo alcune sue notevoli recensioni (a Cantor, Husserl, Cohen), vari saggi sui fondamenti della geometria e una trilogia logica composta dai tre saggi *Der Gedanke – Eine logische Untersuchung* (*Il pensiero – Una ricerca logica*) del 1918; *Die Verneinigung – Eine logische Untersuchung* (*La negazione – Una ricerca logica*) dello stesso anno; *Logische Untersuchungen – Dritter Teil: Gedankengefüge* (*Ricerche logiche – Terza parte: strutture di pensiero*) del 1923.

riguardati come «primitivi», irriducibili: primo fra tutti, ovviamente, lo stesso concetto di numero naturale;

b) *derivare* le «verità» della matematica pura (e in particolare quelle ritenute più «evidenti») a partire da principi meramente logici, e impiegando metodi di ragionamento del tutto esplicitati.

Sullo sfondo di questo programma stanno i dibattiti di cui sopra ci siamo occupati, ma quello che è essenziale per afferrare il senso del nuovo ruolo che Frege assegna alla logica è il riconoscimento di come, per realizzarlo, si ponesse come centrale per Frege una preliminare sistemazione della logica stessa (in quanto teoria dell'inferenza); sistemazione questa che è subordinata, strumentale al problema centrale, quello di assicurare una fondazione certa, incontrovertibile per la matematica. Frege d'altra parte è convinto che la sua proposta in questo senso «... non potrà apparire come una delle tante opinioni, tutte ugualmente giustificabili»; egli ritiene di giungere a una «soluzione definitiva del problema, almeno nei punti fondamentali». E quali per lui siano questi «punti fondamentali» può essere già riconosciuto considerando le varie «polemiche» che Frege condusse, nelle direzioni più diverse, durante la sua attività scientifica e che sottolineano lo stretto legame fra motivazioni filosofiche e matematiche che è caratteristico della sua concezione del ruolo della logica. Lo sfondo filosofico, il punto di riferimento obbligato di queste ricerche rimaneva pur sempre la *Critica*: questo è un fatto che non riguardava solo Frege e portava all'instaurarsi di un abito mentale in base al quale problemi apparentemente specifici della matematica venivano costantemente inseriti in, o confrontati con, un contesto più generale mentre, d'altra parte, si riconosceva che determinati problemi tradizionalmente assegnati all'ambito filosofico non possono risolversi senza l'ausilio della matematica (e abbiamo in questo senso già visto alcuni esempi tipici).

Frege dichiara esplicitamente: «... una ricerca profonda del concetto di numero... costituisce un compito comune alla matematica e alla filosofia»; anzi, proprio in questo riferimento filosofico generale egli individua i «bersagli» delle sue diverse discussioni e polemiche: l'empirismo matematico e lo psicologismo logico di J.S. Mill e di Benno Erdmann; l'«intuizionismo» matematico di Kant; il formalismo prima ingenuo di Thomae e Hankel e suc-

cessivamente più maturo e articolato di Hilbert. Il tenore di queste polemiche, che non risparmiano neppure Dedekind e Cantor, è assai diverso da caso a caso; sicché ci sembra che esponendo molto brevemente il loro contenuto si abbia una prima approssimativa delineazione della stessa posizione di Frege.

4.1 *Frege «contro» predecessori e contemporanei*

Intransigente e irriducibile è la presa di posizione di Frege contro Mill: a suo parere con le premesse di questo autore si giunge in definitiva, circolarmente, a confondere «sempre la pura proposizione aritmetica con le applicazioni che se ne possono fare, le quali sono spesso di ordine fisico e si riferiscono a fatti osservati»; ne può risultare al più un'aritmetica dei «granellini di pepe o dei sassolini», non certo la fondazione di una *scienza* dei numeri. Frege concluderà viceversa che le «proprietà dei numeri derivano dalla loro definizione ed è... presumibile che il metodo induttivo stesso... possa venir giustificato soltanto per mezzo dei teoremi generali dell'aritmetica». Le varie forme di psicologismo hanno d'altra parte precise e decisive responsabilità: «Se malgrado alcuni tentativi dalle due parti, una collaborazione fra matematica e filosofia non è ancora così feconda come sarebbe desiderabile e certo anche possibile, ciò dipende... dal sopravvento preso dai metodi psicologici nella filosofia e dal loro infiltrarsi anche nella logica». Perché, continua Frege, è ben vero che «... può essere utile studiare il flusso di rappresentazioni[33] che accompagnano il pensiero matematico; non si illuda però la psicologia di poter contribuire con ciò alla fondazione dell'aritmetica». In altri termini, la base logica del ragionamento non va mai confusa con le condizioni soggettive interne o esterne al modo particolare con cui viene condotto. Viceversa – e qui interviene l'esplicita professione della sua propria convinzione – «quanto più la matematica deve astenersi da qualsiasi ricorso a considerazioni psicologiche, tanto meno può negare, invece, i suoi rapporti con la logica. Io mi trovo veramente d'accordo con coloro i quali ritengono impossibile tracciare una precisa linea divisoria tra le due».

[33] Frege usa questo termine sempre ed esclusivamente in senso soggettivo.

Per quanto riguarda la confutazione della teorizzazione kantiana dei giudizi aritmetici come sintetici *a priori*, il tono è qui assai più impegnato e si direbbe riguardoso, come deve essere parlando «di un genio cui possiamo guardare solo con riconoscente ammirazione» e della cui teoria si mettono in luce solo «piccole manchevolezze». Il fatto è che la reale e più profonda motivazione filosofica dell'opera di Frege è proprio quella di dare una risposta da una parte alla questione se gli oggetti aritmetici (i numeri) siano afferrabili grazie a un'intuizione interna o sensibile dell'uomo, oppure per altra via; dall'altra «se le leggi dei numeri siano verità analitiche o sintetiche, *a priori* o *a posteriori*». Questo è – per inciso – un tipico esempio per Frege di problemi che pur assegnati tradizionalmente alla filosofia non possono risolversi «senza l'aiuto della matematica».

Circa la prima questione Frege ritiene che Kant sia potuto giungere alla conclusione che il numero è qualcosa di intuitivo e che le *formule* aritmetiche siano di conseguenza indimostrabili e intuitivamente chiare (come gli assiomi e sintetiche) solo perché «ha tenuto conto soltanto dei numeri piccoli». Per quanto invece riguarda il discorso sulla natura delle *leggi* aritmetiche, Frege dichiara che non intende introdurre significati nuovi per i termini analitico, sintetico, *a priori* e *a posteriori*, ma anzi cogliere più compiutamente – ampliandolo – lo stesso pensiero di Kant in proposito; egli afferma che a suo parere le suddette distinzioni non riguardano «il contenuto del giudizio, ma la sua giustificazione», tanto che ove «manca quest'ultima cade... la possibilità stessa di una tale suddivisione». Nel caso dei giudizi matematici, tale giustificazione non può essere che una *dimostrazione* che li riconduca alle verità fondamentali; saremo allora di fronte a una verità analitica quando nel corso di tale dimostrazione «si fa esclusivamente uso delle leggi logiche generali e di qualche definizione precisa», mentre qualificheremo come sintetica una proposizione aritmetica nella cui dimostrazione deve intervenire «qualche verità che risulti non di natura logica generale, ma dipendente da un campo particolare della scienza».

Frege conclude che Kant ha sottovalutato i giudizi analitici, concorrendo così autorevolmente a mantenere in vita la «leggenda» dell'infecondità della logica pura; a suo parere invece in un giudizio analitico sono sì contenute tutte le sue conseguenze, ma «come la pianta nel seme, non come una trave nella casa». Per of-

frire una dimostrazione completa e convincente della sostenuta natura analitica delle leggi aritmetiche, occorre a Frege porsi in grado di condurre dimostrazioni, «lunghe catene di ragionamenti», senza lacune, ossia in modo tale che *ogni* presupposto intervenga esplicitamente nella catena deduttiva e se ne possa quindi valutare la «natura», giusto il canone di analiticità sopra enunciato.

Una terza, caparbia e continua polemica che occupa Frege praticamente durante tutta la sua lunga attività scientifica, è quella che egli conduce contro il formalismo. La polemica si sviluppa sostanzialmente in due fasi. Dapprima egli si impegna contro una concezione formalistica dell'aritmetica intesa, ed effettivamente così presentata da Hankel e Thomae, come una concezione che molto semplicisticamente vede nel *segno* in quanto tale l'oggetto ultimo della ricerca matematica; a questa veduta Frege oppone la sua concezione *contenutistica*, secondo la quale il segno numerico non è che lo strumento per denotare quello che è il vero e proprio oggetto dell'aritmetica, ossia il *numero*. Anche se già in questo contesto Frege avanza argomentazioni che poi saranno portate contro Hilbert, la sostanza della confutazione fregeana consiste nel dimostrare (in varie forme e momenti diversi) come un'aritmetica di tipo formalista riesce a «sostenersi» e a enunciare leggi aritmetiche generali soltanto se, inconsapevolmente ma ineluttabilmente, ricorre a quel «contenuto» del segno numerico del quale di principio dichiara di poter fare completamente a meno. Il tono con cui Frege conduce questa polemica contro l'aritmetica formalista è decisamente presupponente; va detto però che in questo caso le sue critiche «centrano» effettivamente il bersaglio, anche se ciò dipende soprattutto dalla pochezza e dallo scarso sostrato teorico con i quali i suoi oppositori sostenevano la concezione suddetta.

Ben diverso l'impegno richiesto a Frege nella seconda fase di questa polemica, che ora egli conduce direttamente con Hilbert (o successivamente con qualche suo allievo). La posizione hilbertiana è indubbiamente più profonda e articolata, e ben diversa la statura intellettuale di Hilbert rispetto ai primi formalisti: ne viene che nella discussione vengono subito toccate questioni più significative e delicate delle rispettive concezioni dei due contendenti (per Hilbert si veda il prossimo paragrafo). Per Frege, la non contraddittorietà di un *concetto* non garantisce per nulla l'*esistenza* di

oggetti che cadano sotto quel concetto: ché anzi, a suo parere, l'unico modo per dimostrare la non contraddittorietà di un concetto consiste proprio nell'esibire degli oggetti che cadano sotto quel concetto. Nel ruolo particolarmente delicato assegnato da Hilbert alla non contraddittorietà si potrebbe dunque riscontrare una sorta di circolarità. In particolare, continua la critica di Frege, quello di esistenza non è un concetto paragonabile ad esempio a «numero pari»: il primo riguarda (ossia è applicabile, *in* esso cadono solo) *concetti*, mentre il secondo invece riguarda (è applicabile, *sotto* di esso cadono soltanto) *oggetti*. Vedremo nel paragrafo 5 quali siano le ragioni di Hilbert.

Quelle che per Frege sono le ineliminabili deficienze costituzionali del formalismo hilbertiano sono proprio una non osservata distinzione fra oggetti e concetti da una parte e, dall'altra, l'assenza di una *gerarchizzazione* dei concetti. Sta di fatto tuttavia che Frege dimostra di non aver compreso il senso della «definizione implicita» hilbertiana e in realtà questa polemica è più una dimostrazione di reciproca incomprensione che un proficuo scambio di idee. Ciò può essere dovuto a un certo «dogmatismo» da parte di Frege, derivantegli dalla convinzione di possedere la soluzione «definitiva» al problema della fondazione, e alla preoccupazione, da parte di Hilbert, di diffondere nel mondo matematico la sua nuova concezione assiomatica, il che non lo portava a indulgere a polemiche o a controbattere a certe troppo sottili distinzioni non «operative».

Sarebbe in effetti arduo trovare anche un solo contemporaneo di Frege interessato in un modo o nell'altro al problema della fondazione al quale egli non abbia rivolto precise e talora pedanti critiche almeno per quanto riguarda il momento più specificamente «logico» di questo interesse. Così ad esempio avviene con Cantor[34] al quale Frege – pur essendo uno dei pochi ad apprezzare in pieno il valore dell'estensione cantoriana del concetto di numero – rimprovera il residuo psicologistico nella concezione dell'astrazione e in generale una mancanza di rigore logico in tema di defini-

[34] Si noti peraltro che Frege assume le difese di Cantor contro Illigens che, criticando la teoria cantoriana dell'irrazionale, lo aveva classificato fra i matematici formalisti. In questa occasione Frege ribadisce di considerare Cantor un *contenutista* e avanza poi contro di lui, fra l'altro, le obiezioni di cui abbiamo già parlato nel paragrafo 2.

zioni.[35] A Dedekind infine Frege contesta l'ammissibilità del preteso «potere creativo» delle definizioni (probabilmente non avendo compreso a fondo il ruolo di questa «creatività» nell'elaborazione dedekindiana); ma fa in proposito dichiarazioni anche molto più impegnative come ad esempio: «Anche Dedekind è dell'opinione che la teoria dei numeri sia una parte della logica; ma il suo scritto [*Essenza e significato dei numeri*] dà ben pochi elementi per sostenere questa tesi, perché le espressioni da lui usate: "sistema", "una cosa appartiene a un'altra", non sono usuali in logica, né sono riconducibili a qualcosa che sia riconosciuto come logica.» Per noi quest'affermazione è alquanto sorprendente perché oggi riconosciamo che sostanzialmente i «linguaggi» di Dedekind e Frege si equivalgono; ma anche se Frege opponeva come logico, a quello di Dedekind, un linguaggio in termini di «concetti» (invece che di «sistemi» o classi) e «relazioni» (si potrebbe dire: una *logica generale* contro una *teoria* di tipo insiemistico) proprio su questo punto le stesse determinazioni di Frege – come vedremo – non mancavano di notevoli ambiguità.

Come abbiamo già avuto occasione di osservare, le analogie con Dedekind erano molto più accentuate – almeno nella sostanza del piano operativo – di quanto Frege non fosse disposto ad ammettere; in particolare «sullo stesso terreno» essi si muovevano – come aveva riconosciuto Dedekind – relativamente alla giustificazione e al ruolo dell'induzione matematica. «Il presente studio», scrive ad esempio Frege nelle *Grundlagen* del 1884, «mostrerà che anche un ragionamento come quello per passare da n a $n + 1$, in apparenza caratteristico per la matematica, riposa su leggi logiche generali, e non necessita di alcuna legge speciale del pensiero associativo» e il metodo fregeano di caratterizzazione della successio-

[35] Qui ovviamente Frege ha ragione. Abbiamo già avanzato delle perplessità circa la definizione cantoriana di numero cardinale e di tipo d'ordine. Ora è chiaro che nel caso di un insieme con un solo elemento non è possibile «astrarre» nel senso di Cantor dall'ordine; o, pensando invece all'insieme vuoto, non è possibile nemmeno astrarre dalla natura degli elementi. Queste sono le «astrazioni impossibili» che secondo Frege Cantor richiede: in effetti, sulla base della definizione non «esisterebbe» il cardinale 0 né l'ordinale 1. Vedremo come Frege darà queste definizioni. Per quanto invece riguarda il momento psicologistico dell'astrarre nel senso di «trascurare delle proprietà», Frege trasporta la questione a livello logico traducendolo nella sostituzione di una *variabile* a una costante (si veda più avanti la nota 38).

ne dei naturali può essere senza difficoltà reso in termini di catene dedekindiane, ove per elemento base si assuma lo 0 invece dell'1. Le differenze fra i due autori andavano in effetti cercate in altra direzione, sostanzialmente nel concetto stesso di *logica* cui essi si riferivano. Dedekind identificava logica e aritmetica intendendo che la logica fosse *aritmetica*: e in questo senso può considerarsi come l'ultimo anello di quella catena che partendo da Boole giunge a lui attraverso Schröder. Per Dedekind l'aritmetica è *già in se stessa* logica in quanto il concetto di numero non ha un ruolo descrittivo (non costituisce la traduzione concettuale di una realtà specifica) ma un ruolo conoscitivo: consentire l'introduzione di distinzioni che ci fanno capire meglio le proprietà degli oggetti. È il ruolo che dà significato logico al numero, non la sua specifica costruzione. Per Dedekind «logica» è un termine abbastanza indeterminato che si riferisce all'organizzazione del pensiero, alla struttura di fondo degli strumenti conoscitivi. Altra è l'idea di logica di Frege: i principi fondamentali della matematica (geometria esclusa) sono *logici* nel senso che sono leggi logiche, vere sempre; specifici sono i concetti introdotti. È così che l'aritmetica «diventa soltanto una logica altamente sviluppata, ogni teorema aritmetico una legge logica, anche se derivata». Le stesse «applicazioni dell'aritmetica alla spiegazione della natura» sarebbero per Frege «rielaborazioni logiche di fatti osservati», logiche nel senso che riguardano non gli oggetti esterni ma le proposizioni su di essi che sono oggetti logici secondo la definizione di numero che Frege dà e in base alla quale – come vedremo – i numeri divengono proprietà di proprietà. Il problema è quello di *ricostruire* i concetti aritmetici specifici a partire da concetti logici.

Pur se strettamente legata alla concezione stessa del suo programma, la riduzione dell'aritmetica alla logica nel senso di Frege ha almeno una conseguenza generalissima, che trascende i limiti del programma stesso e in particolare è del tutto indipendente dal fatto che esso venga o meno realizzato. Mentre infatti in tutti i tentativi finora visti, il paradigma del procedere deduttivo rigoroso veniva individuato in modo naturale e non problematico nella *logica intuitiva* della matematica (e in questo senso oltre che ovviamente Dedekind si deve tenere presente in particolare anche Cantor), viene ora posta in evidenza da Frege la necessità di indagare sulla natura stessa dell'argomentazione deduttiva, di porre

cioè al centro dell'analisi il concetto stesso di *dimostrazione*: è questo concetto che va rigorosamente reso esplicito in ogni suo passo operativo; vanno chiaramente isolati ed esplicitati le *regole* e i *principi* sui quali la deduzione si fonda, in modo tale che ogni passaggio avvenga secondo quelle regole e ogni proposizione suscettibile di dimostrazione venga ricondotta a quei principi o, il che è lo stesso, da essi derivata.[36]

4.2 *Scansione cronologica e momenti fondamentali della teoria logica di Frege*

Sono già emersi nelle considerazioni precedenti, alcuni motivi fondamentali, alcune «costanti», della concezione fregeana. Vediamo ora di precisare più da vicino gli elementi essenziali di tale teoria, premettendo un'elencazione delle principali opere nelle quali essa viene prendendo corpo, nell'ambito del programma generale di riconduzione della matematica (sempre con l'esclusione della geometria) alla logica. Tale programma è già presente in Frege nel 1879 quando nel volume *Begriffsschrift – Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (*Ideografia – Un linguaggio in formule del pensiero puro a imitazione di quello aritmetico*) presenta un simbolismo adeguato alla completa esplicitazione dei passi e delle ipotesi del processo deduttivo, una logica che realizzi concretamente tale processo, e una teoria generale delle successioni che costituisce già di per sé materiale sufficiente alla giustifica-

[36] Da questo punto di vista, il fatto che il ragionamento matematico abbia caratteristiche che più di ogni altro lo avvicinano a questo nuovo canone di rigore, non ha alcun particolare ruolo decisivo. In effetti, questa è la vera rivoluzione portata nella logica da Frege. In questo senso la sistemazione tecnica della stessa, che pur non perde minimamente la sua importanza decisiva, si presenta tuttavia nella sua giusta luce e refuta già all'origine tutta una serie di critiche (specialmente di parte filosofica) ancora oggi tranquillamente correnti (al solito, in particolare in Italia) circa la pretesa sterile «tecnicità» della logica al confronto con le speculazioni «dense di pensiero» del filosofo tradizionale. Il fatto è che Frege offre un esempio clamoroso di come si possano conciliare e affiancare armonicamente le due esigenze di un discorso significativo condotto con mezzi controllati e rigorosi. Probabilmente neppure Frege presagiva, quando si apprestava a concludere trionfalmente la sua fatica nel 1903, di quali e quanti sviluppi *autonomi* sarebbe stata capace quella logica che seppur fondamentale era da lui riguardata comunque come lo strumento, l'*organon* per eccellenza.

zione «logica» dell'induzione matematica. Più impegnati filosoficamente il già citato saggio del 1884, *Die Grundlagen der Arithmetik. Eine logische-matematische Untersuchung über den Begriff der Zahl* (*I fondamenti dell'aritmetica. Ricerca logico-matematica sul concetto di numero*)[37] e i saggi del periodo 1890-92: *Über das Tragheitsgesetz* (*Sul principio d'inerzia*, 1890), *Über Begriff und Gegenstand* (*Oggetto e concetto*, 1892), *Über Sinn und Bedeutung* (*Senso e significato*, 1892), *Funktion und Begriff* (*Funzione e concetto*, 1892). Nel saggio dell'84, un vero gioiello della letteratura filosofico-scientifica, Frege presenta in maniera non formale tutto il suo programma, motivandolo ampiamente da un punto di vista filosofico e rimanda esplicitamente alla necessità di una trattazione simbolica della parte propriamente esecutiva del programma stesso (che inizierà nel 1893). Nel secondo e nel terzo dei saggi sopra citati, in chiave antipsicologistica tratta della distinzione fra rappresentazione (soggettiva) e concetto (oggettivo) da una parte e fra oggetto (in quanto «saturo», «in sé concluso») e concetto (in quanto «insaturo») dall'altra. Nel quarto saggio presenta una propria originale teoria del significato, nell'ultimo infine dà un'interpretazione funzionale del concetto. La vera e propria esecuzione del programma logicista si avrà nei due volumi *Grundgesetze der Arithmetik* (*Principi dell'aritmetica*, rispettivamente del 1893 e del 1903), bruscamente interrotti dalla comunicazione da parte di Russell dell'antinomia che vanificava (almeno da un punto di vista programmatico) tutto il lavoro di Frege. Occupiamoci ora, anche se brevemente, di alcuni dei temi sopra citati.

È già stata riconosciuta – e da più punti di vista – come essenziale per il programma generale di Frege, la possibilità di condurre dimostrazioni senza «lacune» e del tutto esplicite in ogni loro passo. Orbene, osserva Frege, se si tenta di procedere in questo senso servendosi del linguaggio comune, ci si deve ben presto fermare di fronte ad almeno due difficoltà: da una parte, la complessità sempre crescente delle formulazioni linguistiche ordinarie, in diretta conseguenza appunto dell'analisi minuta dei vari elementi della dimostrazione stessa; dall'altra, la possibilità sempre crescente di fondarsi inavvertitamente su ipotesi intuitivamente ac-

[37] È molto probabile che Frege abbia presentato questo saggio in linguaggio ordinario per reagire alla freddezza con cui era stata accolta l'*Ideografia*, che è a trattazione quasi esclusivamente simbolica.

cettate e non debitamente esplicitate, proprio per associazioni suggerite dalla assuefazione al linguaggio comune. Si impone quindi il ricorso a un simbolismo «asettico», pregnante solo su un piano logico. La scelta di Frege in proposito è molto complessa, elaborata e indubbiamente funzionale ai suoi scopi ma certamente non felice: non è escluso che a tale scelta vada imputata buona parte almeno dell'indifferenza con cui i suoi contemporanei guardarono ai suoi lavori. Ciò non toglie che in un senso l'*Ideografia* segni il vero e proprio atto di nascita della logica matematica quale noi oggi la concepiamo; essa ha aperto la strada a un tipo di applicazione della logica all'indagine della struttura della matematica che trascende gli orizzonti tanto della classica logica formale, quanto dell'algebra della logica di tradizione booleana. È chiaro in che senso l'opera di Frege non sia riassorbibile nella tradizione della logica formale: da un lato perché dà la prima formulazione coerente dell'intergioco fra logica dei termini e logica delle proposizioni e dall'altro in quanto potenzia la logica formale classica in modo assolutamente inedito rispetto alla tradizione.

Più sottile è il diaframma che separa la logica fregeana dall'algebra della logica, anche nei suoi sviluppi dopo Boole. La differenza non è solo tecnica, ma di prospettiva: se gli algebristi miravano ad una teoria delle proposizioni, delle classi, dei termini, ecc., che ponesse in luce l'intelaiatura dei rapporti formali fra questi enti logici, l'intenzione di Frege, come ebbe a scrivere nel 1882 «non era di rappresentare una logica astratta in formule, ma di esprimere un contenuto attraverso segni scritti in modo più preciso e più chiaro di quanto non sia possibile con le parole. Più precisamente quello che volevo creare non era un puro *calculus ratiocinator*, ma una *lingua characterica* nel senso di Leibniz». La differenza è centrale in quanto significa che la scelta degli aspetti formali delle relazioni fra proposizioni, concetti, ecc. da porre in luce, non dovrà essere modellata su una nozione più o meno fantomatica di «struttura logica» che in concreto, come per la maggior parte degli algebristi della logica, si riduceva alla sillogistica classica, ma sui contenuti che la scienza, e in particolar modo la matematica, analizzano. Come scrive Frege «la meta finale dei miei tentativi è una *lingua characterica* destinata innanzitutto alla matematica e non un *calculus* limitato alla logica pura. Il contenuto però deve essere reso più esattamente che nella lingua naturale. Questa infatti la-

scia sempre qualcosa, se pur minima, da indovinare. La composizione delle parole corrisponde solo imperfettamente alla struttura dei concetti... La lingua allude mediante tratti inessenziali o con similitudini a ciò che una ideografia deve esprimere compiutamente».

Dietro l'ideografia fregeana quindi sta tutta un'analisi di nozioni come quello di giudizio, concetto, oggetto, modo di composizione dei pensieri, che rappresentano in sé veri e propri gioielli di indagine logica sul piano filosofico e che mostrano come una delle non ultime ragioni della novità dell'approccio di Frege sia stata quella di porre concretamente alla prova della formulazione matematica nozioni filosofiche tradizionali. Accenneremo più avanti ad alcuni di questi temi; ora invece ci sembra necessario vedere più da vicino l'articolazione della *lingua* in base alla quale Frege costruirà il suo *calculus*. Per Frege il punto di partenza sono le proposizioni che esprimono *contenuti* passibili di *giudizio* e il giudicare, cioè il formulare un giudizio, consiste per lui nell'attribuire a un contenuto uno dei due valori di verità Vero o Falso. È questa nozione di giudizio che porta Frege ad una nuova nozione di calcolo logico che si contrappone a quella degli algebristi; come scriverà nel 1918: «Come la parola "bello" indica l'indirizzo dell'estetica e "buono" quello dell'etica, così la parola "vero" indica l'indirizzo della logica. È ben certo che tutte le scienze si prefiggono la verità, ma la logica se ne occupa in modo diverso. Essa si occupa della verità nello stesso modo in cui la fisica si occupa della gravità o del calore. Scoprire verità è compito di tutte le scienze; alla logica spetta scoprire le leggi dell'esser vero». È appunto questo lo scopo di un calcolo logico, in cui si distingue fra *contenuto giudicabile* e *atto del giudicare*. Allo scopo Frege introduce un *segno di contenuto* — e un *segno di giudizio* |, sicché una scrittura del tipo

$$\vdash \mathscr{A}$$

significa che il contenuto (la proposizione) $\mathscr{A}$ è giudicato vero. Il calcolo logico ci dà assiomi e regole mediante le quali, attraverso il processo di deduzione, tutte le leggi logiche del tipo $\vdash \mathscr{A}$ relative ai contenuti esprimibili nel linguaggio, risultano teoremi.

Il problema allora è come trovare un linguaggio in grado di esprimere contenuti significativi e quali regole di deduzione utiliz-

zare. I due problemi sono connessi in quanto, come afferma Frege «il vero vantaggio si ottiene solo quando il contenuto non viene soltanto indicato, bensì costruito a partire dalle sue parti componenti mediante quegli stessi segni logici impiegati per calcolare». È proprio per questa ragione che il linguaggio di Frege pone sullo stesso piano quello che noi chiameremmo aspetto proposizionale e aspetto predicativo. Boole distingueva tra «primary propositions» che riguardano rapporti fra termini, e quindi gli oggetti della scienza di cui ci si sta occupando e «secondary propositions» che riguardano invece i rapporti fra proposizioni: tra i due aspetti c'è coordinazione, ma non intergioco. Il linguaggio di Frege invece permette di sviluppare sullo stesso piano tanto la logica delle proposizioni che quella dei termini attraverso la distinzione fra proposizioni e funzioni proposizionali,[38] queste ultime intese a esprimere proprietà o relazioni di individui, cosicché tutti i rapporti di inclusione fra classi si possono analizzare in termini di chiusure universali di implicazioni fra funzioni proposizionali. Seguendo quindi una linea analoga a quella di McColl, che subordinava la logica dei termini a quella delle proposizioni così contrapponendosi all'algebra della logica booleana, Frege può correttamente affermare che «la differenza vera consiste nel fatto che io evito una separazione siffatta in due parti, di cui la prima è dedicata ai rapporti fra concetti (*primary propositions*) e la seconda ai rapporti fra giudizi (*secondary propositions*). In Boole le due parti

[38] Una funzione (o forma) proposizionale si ottiene da una proposizione «astraendo» nel senso di Frege da una o più costanti che in essa compaiono e ottenendo così qualcosa di «insaturo». Ad esempio, astrarre da «Antonio» nella proposizione «Antonio ama Maria» significa per Frege sostituire ad «Antonio» una variabile, poniamo x, ottenendo così «x ama Maria» che è appunto una *funzione proposizionale*, la cui insaturazione è denotata proprio dalla presenza della variabile. Tale funzione può essere «saturata» ponendo al posto della x un nome di individuo qualsiasi. È chiaro che viceversa l'operazione di «astrazione» può continuare, ottenendo nel nostro esempio, con una successiva applicazione, il concetto, la forma proposizionale «x ama y» che ora può essere saturata con due nomi di individui, ecc. I *quantificatori* sono simboli per le espressioni «per tutti» (quantificatore universale) o «esiste» (quantificatore esistenziale) mediante i quali è appunto possibile rendere linguisticamente la generalità o l'esistenza. Questa introduzione è forse l'apporto fondamentale di Frege al costituirsi di una «logica dell'inferenza», il punto cruciale di confronto col sistema booleano. Frege in effetti assume un simbolo per il solo quantificatore universale ed esprime poi l'esistenziale come negazione della generalità di una negazione, ossia come «non per tutti non...».

procedono una accanto all'altra, sicché una è come l'immagine speculare dell'altra, ma, proprio per questo, non si trova in un rapporto organico con l'altra».

La soluzione di Frege consiste, esprimendoci in termini odierni nella costruzione di un linguaggio del primo ordine con identità[39] i cui simboli logici sono $\rightarrow$, $\neg$, $\forall$ ed $=$, simboli per variabili individuali e predicati di ogni arietà finita (quindi binari, ternari, ecc.) e lettere proposizionali. Le regole di formazioni sono quelle usuali (le vedremo più avanti, a proposito della sistemazione di Hilbert e Ackermann) e riflettono l'interpretazione intuitiva della composizione di proposizioni più complesse a partire da proposizioni più semplici mediante l'impiego di connettivi e quantificatori. Quello che vogliamo sottolineare subito è che in questo modo la struttura grafica delle formule riflette la «forma logica» della connessione fra contenuti, così come per Frege (come già per Leibniz, del resto) deve fare ogni linguaggio ideografico.

Quanto al calcolo logico, esso si basa sui seguenti assiomi

$$\vdash p \rightarrow (q \rightarrow p)$$

$$\vdash (p \rightarrow (q \rightarrow r)) \rightarrow ((p \rightarrow q) \rightarrow (p \rightarrow r))$$

$$\vdash (p \rightarrow (q \rightarrow r)) \rightarrow (q \rightarrow (p \rightarrow r))$$

$$\vdash (p \rightarrow q) \rightarrow (\neg q \rightarrow \neg p)$$

$$\vdash \neg \neg p \rightarrow p$$

$$\vdash p \rightarrow \neg \neg p$$

$$\vdash x = y \rightarrow (P(x) \rightarrow P(y))$$

$$\vdash x = x$$

$$\vdash \forall x P(x) \rightarrow P(y),$$

e sulla regola del *modus ponens*, esplicitamente formulata come tale:

$$\vdash \mathscr{A}, \vdash \mathscr{A} \rightarrow \mathscr{B} \,/\, \vdash \mathscr{B},$$

[39] Almeno per quanto riguarda la struttura assiomatica del calcolo. Nell'ultimo capitolo dell'*Ideografia* Frege non esita tuttavia a quantificare anche su variabili funzionali (o predicative, diremmo noi oggi) usando così di fatto quello che noi chiameremo una logica del secondo ordine.

congiunta con una regola di sostituzione che però Frege usa di fatto ma non esplicita come tale.

Come detto abbiamo qui formulato gli assiomi (in generale il linguaggio) di Frege in termini moderni, ma il lettore può farsi un'idea dell'effettiva simbolizzazione bidimensionale fregeana dalla figura riportata a pag. 349.

Momento centrale della teorizzazione di Frege è la sua costante preoccupazione di tenere separato e distinto in modo netto e rigido il concetto dalla rappresentazione da un lato e dall'oggetto dall'altro. La prima distinzione avviene sostanzialmente, come si diceva, in chiave antipsicologistica: il concetto è qualcosa di oggettivo e come tale va assegnato alla logica (che ne richiede peraltro l'esatta delimitazione); esso è per sua natura astorico. Un concetto rigorosamente delimitato non ha storia: più appropriato è, semmai, parlare di «storia dei tentativi per comprendere un concetto». La rappresentazione, viceversa, è essenzialmente un fatto soggettivo, «ha bisogno di un portatore»: come tale va «utilmente attribuita alla psicologia».

La seconda distinzione, quella fra concetto e oggetto, è invece di tipo prettamente logico: è proprio del concetto lo specifico carattere di «insaturazione», di predicatività, di fronte all'essenziale natura «satura», di ente «completo» dell'oggetto. A livello simbolico, a un oggetto corrisponde un *nome proprio*, a un concetto un *nome di funzione* (in questo senso noi diremmo oggi una «lettera predicativa»). In *Funzione e concetto* Frege in effetti identifica i concetti con particolari funzioni, precisamente con quelle funzioni il cui valore è sempre, per qualunque argomento, un *valore di verità* e successivamente, nel primo volume dei *Principi*, stabilisce una gerarchizzazione in *gradi* e *tipi* delle funzioni (e quindi in particolare dei concetti) basata sulla natura e sul numero degli argomenti che possono «saturarle» (ossia che possono essere sostituite alle variabili contenute nella funzione). A noi non interessa qui approfondire la questione, il che tra l'altro appesantirebbe di molto l'esposizione. Ci limitiamo a osservare che proprio a causa della rigida distinzione fra concetti e oggetti, questi ultimi non solo sono del tutto estranei alla citata gerarchia, ma anzi vengono assunti senza distinzione di tipo nel senso che un dato oggetto, qualunque sia la sua natura, la sua origine, il suo *status* ontologico, è *da un punto di vista logico* del tutto indistinguibile da un altro. Questo fatto, come

vedremo, avrà particolari ripercussioni se applicato alle classi.[40]

Se la distinzione fra oggetto, concetto e rappresentazione (in particolare la prima) è uno dei temi più caratteristici della logica di Frege, la teoria del significato che egli propone in *Senso e significato* è indubbiamente uno degli apporti fondamentali che egli diede alla speculazione logica in generale. Si tratta brevemente di una proposta di soluzione al problema di associare ad ogni nome proprio, predicato, enunciato (in generale, a ogni elemento linguistico «significante») un significato. Allo scopo Frege distingue due piani di significatività, quello che in terminologia moderna viene detto della connotazione o intensione (o del senso) e quello della denotazione o estensione (o del significato). Il piano dell'estensione si pone a livello oggettuale mentre quello della connotazione è intermedio fra questo e il piano puramente soggettivo della rappresentazione. A una data intensione corrisponde un'unica estensione; viceversa, una data estensione può essere individuata da più intensioni diverse. Ad esempio Socrate è l'estensione dei due termini «il maestro di Platone» e «il marito di Santippe» che hanno contenuto concettuale diverso, diversa intensione, in quanto rappresentano due prospettive distinte lungo le quali possiamo fare riferimento a Socrate. Un'intensione è, in sostanza, un *modo* di denotare un'estensione.

Ogni simbolo linguistico significante, semplice o complesso, *esprime* una intensione e *denota* una estensione. Lasciando da parte il dibattuto problema delle intensioni, limitiamoci a precisare che per Frege la denotazione di un *nome proprio* è un *oggetto*; quella di un *predicato* è una *funzione intesa estensionalmente*;[41] infine la denotazione di un *enunciato* è uno dei due *valori di verità*: Vero o Falso. Nell'ambito di una decisa convinzione platonica, Frege considera il Vero e il Falso due *oggetti*: ne viene che ogni enunciato è un nome proprio, e precisamente il nome appunto di uno dei due valori di verità, fornendo un ulteriore esempio di come espressioni con

[40] Proprio in questo la gerarchia di Frege si differenzia dalla analoga gerarchia definita da Russell nella sua *teoria dei tipi*. Lo stesso Russell riconosceva nell'idea di Frege una precisa anticipazione della propria teoria.

[41] Forzando un po' il linguaggio di Frege possiamo dire che a un predicato viene assegnata come estensione l'insieme degli elementi che godono della data proprietà (se il predicato è a un posto) o fra i quali sussiste la data relazione (se il predicato è a più posti).

uguale estensione (le proposizioni vere) possono avere intensione, contenuto concettuale, diverso, in quanto chiaramente non tutte le proposizioni vere «dicono la stessa cosa», esprimono lo stesso pensiero.

Non possiamo qui seguire l'evoluzione del simbolismo e della stessa logica di Frege nel corso del suo sviluppo teoretico. Fondamentale però è il passaggio dal sistema della *Begriffsschrift* a quello dei *Grundgesetze*, dove Frege si porrà come esplicito obiettivo – come vedremo – di definire il concetto di numero naturale e di dimostrare l'esistenza nell'universo degli oggetti logici della estensione di questo concetto. Nei *Grundgesetze* la logica, che ora ha un obiettivo più impegnativo e deve essere tanto potente da dimostrare principi esistenziali, diviene una logica, diremmo noi oggi, del secondo ordine. Ciò significa, al solito, che si possono quantificare variabili predicative e che si ammettono nello stesso apparato linguistico espressioni per denotare i *decorsi di valori* di funzioni proposizionali, ossia quello che potremmo chiamare l'oggetto corrispondente ad ogni dato concetto (la sua estensione). In questo modo già nell'apparato linguistico è assunto un principio esistenziale che ci consente di «scrivere» dato ogni concetto – nella notazione di Frege $f(x)$ – la corrispondente «estensione», o meglio il *nome* di tale estensione, in notazione fregeana $\grave{\varepsilon} f(\varepsilon)$. Sul piano deduttivo il sistema che ne emerge – la Grande Logica fregeana entro cui ricostruire l'aritmetica – sarà corredato da un numero molto maggiore di regole di inferenza e da sei assiomi (che svolgono per il linguaggio più ricco la funzione dei precedenti assiomi dell'*Ideografia*) il quinto dei quali riguarda esplicitamente il rapporto fra concetti e loro decorsi di valori stabilendo che

$$\forall x\ (f(x) = g(x)) \longleftrightarrow \grave{\varepsilon} f(\varepsilon) = \grave{\alpha}\ g(\alpha)$$

cioè: due decorsi di valori (estensioni) sono uguali come oggetti se e solo se sotto i corrispondenti concetti cadono gli stessi oggetti. Questo ci permette di chiarire un punto cui sopra abbiamo accennato senza soffermarci: quando abbiamo detto che fra i simboli logici della *Ideografia* figurava anche l'identità (=) non abbiamo sottolineato come ciò significasse una rottura rispetto alla pratica matematica corrente, nella quale l'identità non è logica, ma «regionale», ossia viene definita di volta in volta a seconda del dominio considerato, per mezzo di criteri in generale diversi. Per Frege

l'identità è una sola, definita una volta per tutte su *tutti* gli oggetti: il dominio della Grande Logica è ontologicamente esaustivo, contiene tutti gli oggetti pensabili.

Begriffe fällt, nicht aber unter den zweiten.

Begriffsschriftlich können wir das so ableiten:

$$\nu \quad \vdash \overset{\mathfrak{g}}{\smile} \; \mathfrak{g}(a) \;;\; M_\beta(-\mathfrak{g}(\beta)) = a \;;\; M_\beta\left(\overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(\beta) \;;\; M_\beta(-\mathfrak{g}(\beta)) = \beta\right) = a$$

(IIIa): ————————————

$$\vdash f(a) \;;\; M_\beta\left(\overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(\beta) \;;\; M_\beta(-\mathfrak{g}(\beta)) = \beta\right) = a \;;\; f(a) = \overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(a) \;;\; M_\beta(-\mathfrak{g}(\beta)) = a \qquad (\xi$$

(IIb):: - - - - - - - - - - - - - -

$$\vdash f(a) \;;\; M_\beta\left(\overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(\beta) \;;\; M_\beta(-\mathfrak{g}(\beta)) = \beta\right) = a \;;\; M_\beta(-f(\beta)) = M_\beta\left(\overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(\beta) \;;\; M_\beta(-\mathfrak{g}(\beta)) = \beta\right) \;;\; \overset{\mathfrak{F}}{\smile}\; \mathfrak{F}(a) = \overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(a) \;;\; M_\beta(-\mathfrak{g}(\beta)) = a \;;\; M_\beta(-\mathfrak{F}(\beta)) = M_\beta\left(\overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(\beta) \;;\; M_\beta(-\mathfrak{g}(\beta)) = \beta\right) \qquad (o$$

(IIb, IIIa):: = = = = = = = = = = = = = = = =

$$\vdash f(a) \;;\; M_\beta(-f(\beta)) = a \;;\; M_\beta\left(\overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(\beta) \;;\; M_\beta(-\mathfrak{g}(\beta)) = \beta\right) = a \;;\; \overset{\mathfrak{G}\;\mathfrak{F}}{\smile}\; \mathfrak{F}(a) = \mathfrak{G}(a) \;;\; M_\beta(-\mathfrak{F}(\beta)) = M_\beta(-\mathfrak{G}(\beta)) \qquad (\pi$$

$$\smile$$

$$\vdash \overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(a) \;;\; M_\beta(-\mathfrak{g}(\beta)) = a \;;\; M_\beta\left(\overset{\mathfrak{g}}{\smile}\; \mathfrak{g}(\beta) \;;\; M_\beta(-\mathfrak{g}(\beta)) = \beta\right) = a \;;\; \overset{\mathfrak{G}\;\mathfrak{F}}{\smile}\; \mathfrak{F}(a) = \mathfrak{G}(a) \;;\; M_\beta(-\mathfrak{F}(\beta)) = M_\beta(-\mathfrak{G}(\beta)) \qquad (\varrho$$

(μ): - - - - - - - - - - - - - - -

Una pagina dai *Principi* di Frege. La cavità con lettere gotiche nel segno di contenuto è il segno per il *quantificatore universale*; i diversi segni fra una formula e la successiva indicano le diverse regole applicate in quel passaggio.

Proprio per questa sua esaustività è secondo Frege legittima l'assunzione, che si può facilmente provare in base all'assioma dei *Grundgesetze*, di quel *principio di comprensione* secondo il quale per ogni concetto esiste la corrispondente estensione. Si noti che si ha qui per la prima volta una considerazione *esplicita* di tale cruciale principio (già utilizzato, come si ricorderà, anche da Cantor) e che fornisce a Frege la base per considerare come oggetti dell'universo logico gli enti matematici che vengono introdotti.

Un'ultima osservazione ci sembra opportuna in questo contesto. Si era detto che la logica di Frege si presentava come una logica intensionale, relativa cioè a concetti, non a estensioni di concetti e nell'*Ideografia* le cose stanno effettivamente così. Ma avevamo anche detto che la questione in Frege è ambigua, perché talora egli afferma di poter trattare invece che con «concetti» con «estensioni di concetti» senza peraltro mai chiarire che cosa intenda effettivamente con la possibilità di questa sostituzione. È ben vero che Frege afferma esplicitamente, in proposito: «Si può forse avere l'impressione che nel conflitto fra logici estensionali e intensionali io mi ponga dalla parte di questi ultimi. In effetti io ritengo che il concetto è logicamente antecedente alla sua estensione e riguardo come futile il tentativo di fondare l'estensione di un concetto come classe non sul concetto stesso, ma sulle singole cose (sugli individui)»; ma questa è tutto sommato una dichiarazione di principio che non chiarisce in modo particolare la questione.

Va invece detto che la logica dei *Principi* si annette una dimensione estensionale del tutto assente nella logica della *Ideografia*: e ciò dipende in modo diretto proprio dall'analisi, compiuta da Frege con la sua teoria del significato, del «contenuto concettuale» di una proposizione in *senso* della proposizione stessa e sua *denotazione*, o valore di verità. Nei *Principi*, Frege usa locuzioni quali ad esempio l'α-concetto A e simili per indicare un concetto la cui estensione sia la classe α. Se ne può concludere che pur mantenendo la priorità logica al concetto (si ricordino le osservazioni a proposito di Dedekind e la precedente dichiarazione di Frege) è semplicemente più «comodo» e meno complicato parlare in termini di estensioni di concetti, di classi. L'iniziale caratteristica intensionale della logica di Frege non viene cioè sostituita, nella più matura sistemazione dei *Principi*, da una concezione estensionale,

bensì semplicemente affiancata da questa. Vogliamo osservare, per finire, che la questione è per così dire indifferente rispetto all'insorgere dell'antinomia di Russell che si presenta tanto in termini di classi (estensioni concettuali) quanto in termini di predicati (concetti).

4.3 *La definizione di numero naturale e la caratterizzazione della successione numerica*

La definizione fregeana del concetto di numero cardinale coincide nella sostanza con quella di Cantor tanto nel senso che entrambi giungono a individuare lo stesso concetto (o, in termini di classe, la stessa classe di elementi) quanto nel senso che entrambi si servono per così dire dello stesso «materiale»: le classi e una relazione fra di esse. Cantor, abbiamo visto, chiama *equivalenza* questa relazione, Frege invece la chiama *equinumerosità*; ma al di là della diversa terminologia si tratta della stessa relazione: in entrambi i casi infatti tale relazione sussiste fra due classi quando è possibile stabilire fra di esse una corrispondenza biunivoca.[42] Ben diversa è tuttavia la costruzione, se così si può dire, delle due definizioni dal punto di vista della correttezza logica. Come abbiamo visto, Cantor giungeva dapprima al concetto generale di numero cardinale o potenza tramite una «astrazione» di marca psicologista perlomeno problematica che non consentiva di capire che tipo di oggetti fossero i numeri così definiti e quindi faceva intervenire la relazione di equivalenza fra classi come *criterio* di uguaglianza fra potenze; si può dire che Frege compia esattamente il cammino inverso, ossia prima riscontra nelle proprietà della relazione di equinumerosità «i caratteri di un'uguaglianza» (e vedremo subito in che senso) quindi sulla base di una rigorosa sistemazione della *definizione per astrazione* ne trae il concetto di numero cardinale. Sarà quindi opportuno, prima di presentare la definizione fregeana, ve-

[42] Analogamente, fra concetti: due concetti sono equinumerosi quando è possibile stabilire una corrispondenza biunivoca fra le rispettive estensioni, ossia fra gli oggetti che cadono sotto l'uno e gli oggetti che cadono sotto l'altro; vista questa possibilità di immediata traduzione e tenuto conto della doppia dimensione estensionale-intensionale della logica di Frege, adotteremo in generale, per comodità, il linguaggio in termini di classi, per noi oggi più usuale.

dere in che cosa consiste una definizione per astrazione (il che noi faremo in termini moderni, senza peraltro minimamente tradire lo spirito almeno, se non la lettera, della discussione di Frege a questo riguardo).

Supponiamo data una classe A di elementi $x, y, z, \ldots$ qualsiasi e supponiamo che fra questi elementi sia definita una relazione R di equivalenza, tale cioè che sia simmetrica, transitiva e riflessiva nel senso che comunque si prendano gli elementi x, y, z di A si abbia:

1) xRx

2) Se xRy allora yRx

3) Se xRy e yRz allora xRz.

Consideriamo ora, per ogni elemento di A, la classe $\alpha(x)$ di tutti gli elementi di A che stanno nella relazione R con x; la classe $\alpha(y)$ di tutti gli elementi di A che stanno nella relazione R con y; la classe $\alpha(z), \ldots$ e così via. Sulla base delle proprietà 1), 2) e 3) si dimostra facilmente che la relazione R determina una ripartizione (che tecnicamente viene detta *partizione*) degli elementi di A in classi *esaustive e disgiunte*[43] $\alpha(x)$, $\alpha(y)$, $a(z)$, ecc. tali che ogni classe contiene, con un dato elemento x di A, ogni altro elemento di A che stia con x nella relazione R e tali inoltre che ognuna di queste classi è *univocamente* determinata da ognuno dei suoi elementi (o, in altri termini, può essere «rappresentata» da uno qualsiasi dei suoi elementi). Orbene, si dice che la classe A_R i cui elementi sono le *classi* $\alpha(x)$, $\alpha(y)$, $\alpha(z), \ldots$ di elementi di A, è stata ottenuta da A mediante una *definizione per astrazione rispetto alla relazione* R. Tutto ciò significa sostanzialmente che abbiamo *identificato parzialmente* gli elementi di A nel senso che, rispetto alla relazione dalla quale si astrae, ogni elemento di una classe è sostituibile in ogni (opportuno) contesto da un altro qualsiasi elemento della stessa classe *salva veritate*: in altri termini ogni classe della partizione contiene elementi *uguali rispetto a* R o quozientando rispetto a R (in questo senso dicevamo prima che Frege aveva colto i caratteri di un'uguaglianza nella particolare relazione da lui considerata).

[43] Ossia ogni elementi di A appartiene almeno ad una di queste classi (esaustività) ed al più ad una di esse (disgiunzione), vale a dire appartiene a una e una sola classe.

Ciò detto, veniamo al nostro caso specifico: si tratta di definire il concetto di numero nel senso di Frege. In questo caso la classe A è la classe di tutte le classi (indichiamola con V) e la relazione R è proprio la relazione di equinumerosità fra classi, che si dimostra immediatamente e facilmente essere riflessiva, simmetrica e transitiva. Tale relazione suddividerà allora la classe V di tutte le classi secondo una partizione in *classi di classi*: ognuna di queste classi di classi sarà caratterizzata, giusto quanto detto sopra, dal contenere, con una data classe α di V, ogni altra classe equinumerosa con α. Orbene Frege identifica il numero cardinale di una classe α con la *classe di tutte le classi ad essa equinumerose.* Risulta allora che ogni numero cardinale è una classe di classi; certamente non sarà vero il viceversa, vale a dire non ogni classe di classi sarà un cardinale. Il concetto generale di numero cardinale (la classe dei numeri cardinali) si otterrà ponendo che un numero cardinale è una classe di classi che è il numero di una data classe (è la classe di tutte le classi di classi che sono numeri cardinali di una qualche classe). Si badi bene che questa definizione non è circolare come l'enunciazione verbale potrebbe far credere: infatti Frege dapprima definisce il *cardinale di una classe*; sulla base di questo concetto ottenuto correttamente per astrazione come abbiamo visto, definisce il concetto più generale di *cardinale* (senza ulteriori qualificazioni).

Si pensi, per togliere ogni dubbio, di voler definire ad esempio la classe dei mariti: prima di tutto si vedrà di definire cosa significa essere marito di qualcuno quindi si potrà passare a definire la classe dei mariti come costituita da tutti coloro che sono marito di qualcuno; in questo procedimento non è insita ovviamente nessuna circolarità, ma semplicemente nei due casi si definiscono concetti (classi) diversi.

Frege ha così definito il concetto (la classe) di numero cardinale che, ce lo ha mostrato Cantor, è ben più ampio del concetto di numero cardinale finito, o numero naturale. Per isolare nella classe dei cardinali la sottoclasse dei naturali, Frege si serve ora della sua teoria generale delle successioni che sostanzialmente altro non è che una teoria dell'induzione generalizzata. È in questo contesto appunto che il discorso di Frege e quello di Dedekind coincidono nella sostanza, anche se i due autori fanno uso di terminologie diverse. Non crediamo opportuno scendere qui in particolari. Ci li-

mitiamo quindi a pochi cenni riassuntivi. Già nell'*Ideografia*, si ricorderà, Frege aveva dato un abbozzo consistente di una sua teoria generale delle successioni. Orbene è il concetto generale di «seguire in una successione» che dà a Frege la chiave dell'induzione, che gli permette di «... ricondurre a leggi logiche generali il processo... che fa concludere da n a $n + 1$». Definito infatti il numero 0 come la classe di tutte le classi vuote (come il numero spettante a un qualunque concetto contraddittorio), Frege riesce, tramite appunto il concetto generale del seguire in una successione, a definire il concetto di successore immediato di un numero sostanzialmente come segue: il successore del numero di una classe α è il numero della classe formata da α assieme con un elemento x che non appartiene ad α, o più suggestivamente, in termini di concetto: «Il numero spettante al concetto "appartenente alla successione dei numeri naturali che termina con d" segue immediatamente d nella successione dei numeri naturali»; assume cioè come successore del numero n il numero della classe formata da tutti i numeri naturali che precedono o sono uguali a n, a partire dallo 0: $\{0, 1, 2, 3, ..., n\}$. Frege dimostra quindi che la successione dei naturali è infinita e il concetto generale di numero naturale viene fatto semplicemente coincidere col concetto di appartenenza alla successione dei numeri naturali che inizia con lo 0 (si ricordi l'osservazione precedente). In altri termini, un numero naturale è «qualunque oggetto» possa essere raggiunto a partire dallo 0 applicando un numero finito di volte l'operazione di successore.[44]

L'analogia col procedimento dedekindiano salta agli occhi: la differenza centrale è però che Frege *definisce* i termini primitivi assunti da Peano nella sua assiomatizzazione e *dimostra* i cinque assiomi peaniani (e quindi le quattro proposizioni di Dedekind) cosa che Dedekind non fa. Per Dedekind l'essenziale era trovare le proprietà *strutturali* che permettono di procedere in modo induttivo e porre in luce – cosa che Frege non fa – la natura del procedimento di definizione per ricorsione. Per Frege invece si tratta di

[44] In questa descrizione abbiamo in effetti un po' «tradito» Frege; ma è oltremodo difficile dare un'idea precisa di quanto sottile e minuziosa sia l'analisi che egli conduce senza scendere in dettagli tecnici che ci paiono fuori luogo in questa sede.

ottenere il concetto di naturale a partire da oggetti che vanno costruiti utilizzando solo materiale logico, proprietà, relazioni, decorsi di valori, ecc. e dimostrare con la *pura logica* che valgono per gli oggetti così costruiti quelle proprietà che caratterizzano i naturali e che egli identifica proprio mediante le stesse proprietà generali usate da Dedekind. Questi enti (i naturali) esistono nell'universo in forza del principio di comprensione che ha un significato logico e non in forza di poteri «psicologici» come in Dedekind. Al fondo di questa differenza c'è una diversa concezione del ruolo dei naturali: per Dedekind essi sono essenzialmente una notazione per indicare iterazioni o comunque definire una struttura ordinale, il loro ruolo è di garantire la possibilità di ordinamento di insiemi di oggetti. Per Frege invece i naturali sono essenzialmente *cardinali*, che indicano la quantità di elementi di una collezione finita: la definizione è sostanzialmente quella di Cantor, posta in veste logicamente ineccepibile e l'isolamento dei naturali come cardinali *finiti* avviene attraverso l'imposizione della validità dello schema di induzione. Ragionamento diverso in Dedekind, per cui la distinzione fra finito e infinito non ha a che fare con la definizione dei naturali (e anzi più tardi si dimostrerà che le due definizioni sono equivalenti solo tramite l'assioma di scelta) ma con l'*esistenza* di un insieme infinito da cui quella dell'insieme dei naturali si può derivare.

Mentre però Dedekind si convince dell'esistenza di almeno un insieme infinito con un ragionamento in ultima istanza di tipo psicologistico, che risale a Bolzano, Frege è consapevole del fatto che allo scopo occorre utilizzare un principio generale per l'esistenza di collezioni: il principio di comprensione, appunto. Ma sarà proprio la indiscriminata generalità di questo principio a far vacillare il sogno fregeano.

4.4 *L'antinomia di Russell*

«A uno scrittore di scienza ben poco può giungere più sgradito del fatto che, dopo completato un lavoro, venga scosso uno dei fondamenti della sua costruzione.» Così Frege dà inizio all'appendice aggiunta al secondo volume dei *Principi* per dare pubblica comunicazione al mondo scientifico del fallimento del suo program-

ma logicista. «Sono stato messo in questa situazione,» prosegue Frege, «da una lettera del signor Bertrand Russell quando la stampa di questo volume stava per essere finita.» Proprio quando Frege stava portando concretamente a termine la riconduzione dell'aritmetica alla logica, il 16 giugno 1902 Russell gli indirizzava infatti una breve lettera nella quale lo informava di un'antinomia derivabile nel sistema logico dei *Principi*. L'antinomia in questione, forse la più celebre di tutte le antinomie antiche e moderne, o comunque senz'altro la più nota almeno per i riflessi immediati che ha avuto, si può brevemente descrivere come segue. In base al principio di comprensione a ogni proprietà corrisponde una classe, un insieme ben determinato, il quale a sua volta può essere considerato come un elemento passibile di eventuali predicazioni, è cioè, in termini fregeani, un oggetto, un ente saturo. Orbene, consideriamo la proprietà «non essere elemento di se stesso»; sulla base di quel principio ad essa corrisponderà la classe di tutte e sole quelle classi che non contengono se stesse come elemento; indichiamo tale classe con R (da Russell) e chiediamoci: R appartiene o no a se stessa, è cioè elemento di se stessa oppure no? Supponiamo che R appartenga a se stessa: allora, in quanto classe di tutte le classi che *non* si appartengono, dovrà non appartenere a se stessa. Supponiamo allora che R non appartenga a se stessa: in questo caso, essa dovrà però appartenere a se stessa, in quanto per definizione è proprio la classe di tutte le classi che non si appartengono. Entrambe le ipotesi portano quindi a una contraddizione, siamo cioè di fronte a un'antinomia. Il ragionamento precedente può essere espresso chiaramente e sinteticamente in forma simbolica come segue. Per ogni classe X, R è definita ponendo

$$X \in R \leftrightarrow X \notin X.$$ [45]

Ponendo in questa $X = R$, si ha subito

$$R \in R \leftrightarrow R \notin R$$

ossia un'equivalenza contraddittoria.

[45] La formula in questione va letta come segue: per ogni classe X, X appartiene a R se e solo se X non appartiene a se stessa, definisce cioè R come la classe di tutte le classi che non si appartengono.

Frege fu comprensibilmente annichilito da questa comunicazione, ma non solamente su un piano «personale»; egli rileva infatti «che tutti coloro che nelle loro dimostrazioni hanno fatto uso di estensioni concettuali, classi, insiemi, sono nella mia stessa situazione. *Qui non è in causa il mio metodo di fondazione in particolare, ma la possibilità di una fondazione logica dell'aritmetica*» [corsivo nostro]. Dei sei assiomi che Frege aveva posto alla base del suo sistema dei *Principi*, l'antinomia colpisce il quinto, che, come sappiamo afferma che se sotto due concetti cadono gli stessi oggetti allora i due concetti hanno estensioni uguali (Va) e *viceversa* se due concetti hanno estensioni uguali allora sotto di essi cadono gli stessi oggetti (Vb).

Precisamente l'antinomia di Russell mette in crisi la parte Vb di tale assioma (e quindi ovviamente tutto l'assioma) e questo è abbastanza chiaro e comprensibile. Infatti, per stabilire la «verità» di Va non è necessario possedere una precisa determinazione del concetto «estensione di un concetto» che anzi può pensarsi proprio definito per suo mezzo; in Vb al contrario, è necessario avere una rigorosa determinazione proprio del concetto «estensione di un concetto» per poter passare a stabilire che esattamente gli stessi oggetti cadono sotto le estensioni di cui si fa parola nell'assioma. Ora l'antinomia di Russell metteva in crisi proprio il concetto di estensione concettuale, perché faceva vedere che l'assumere, col principio di comprensione, l'esistenza e l'«oggettualità» di tale estensione relativamente a *ogni* concetto era un'assunzione contraddittoria.

Della stessa natura erano le difficoltà in cui si era trovato Cantor con l'antinomia di Burali Forti o con quella del massimo numero cardinale e la soluzione cantoriana veniva proprio, implicitamente, a significare una limitazione del principio di comprensione. Quando infatti Cantor giungeva alla conclusione che era necessario distinguere fra «molteplicità incoerenti» e «molteplicità coerenti» egli riconosceva la necessità di distinguere fra due tipi di estensioni concettuali: quelle che potevano considerarsi come una «unità», come un oggetto in termini fregeani, delle quali cioè potevano ulteriormente predicarsi delle proprietà, e quelle invece che esistevano soltanto come molteplicità, che cioè non potevano considerarsi come «insieme» dei loro elementi, ma erano troppo «sfumate», o troppo «grandi» per poter essere considerate come

fatto unitario, come oggetto. È ovvio che nella rigida e indifferenziata determinazione fregeana degli oggetti non poteva trovar posto un tentativo di soluzione dell'antinomia che operasse una tale distinzione fra quei particolari oggetti che erano le estensioni concettuali o classi. Esclusa quindi la possibilità di una soluzione alla Cantor, Frege opera un tentativo che consiste nel mantenere inalterata la validità di Vb solo escludendo dai possibili oggetti che cadono sotto i dati concetti le estensioni stesse dei concetti. Vb diventa cioè: se due concetti hanno estensioni uguali, allora sotto di essi cadono gli stessi oggetti che siano diversi dall'estensione dei due concetti.

Questa soluzione non soddisfece in effetti lo stesso Frege, e fu addirittura a sua volta dimostrata contraddittoria nel 1936. Frege comunque non tornò mai più esplicitamente sul problema dell'antinomia o delle sue soluzioni. Alla fine dell'appendice ai *Principi*, egli aveva indicato, come compito dei fondamenti dell'aritmetica quello di rispondere alla questione di come debbano intendersi gli oggetti logici, e in particolare i numeri, e di stabilire a che titolo siamo autorizzati a riconoscere i numeri come oggetti. Proprio nel periodo in cui, grazie a Russell, la sua opera cominciava a essere conosciuta e apprezzata in tutto il suo valore, Frege si ritira dalla ricerca logica attiva, quale veniva configurandosi in quel periodo, in gran parte proprio a causa dell'antinomia scoperta nel suo sistema: egli lasciò il problema fondamentale di cui sopra alle nuove generazioni e si dedicò alla trattazione di argomenti più filosoficamente impegnati in senso tradizionale.[46]

5. L'ASSIOMATICA MODERNA: DAVID HILBERT

In certo senso «simmetrico» all'intervento di Frege è quello di Hilbert agli inizi del secolo. La problematica relativa al sorgere delle geometrie non euclidee, lo sviluppo dell'algebra e il conseguente costituirsi dell'algebra della logica, gli stessi argomenti finora trattati in questo capitolo, ci hanno già portato in modo na-

[46] Tali articoli sono stati da noi ricordati alla fine della nota 32 a pag. 332.

turale ad accennare alle vedute di David Hilbert[47] come punto d'arrivo sistematico e fecondo di diversi filoni di sviluppo che avevano mutato lo *status* della ricerca matematica e avevano posto in nuova luce la questione dei suoi rapporti con la filosofia. Le considerazioni che seguono a proposito dell'opera di Hilbert ci offrono così la possibilità di tracciare le prime linee di un bilancio del periodo trattato in questo capitolo. Ci interesseremo sostanzialmente di un solo lavoro di Hilbert, le *Grundlagen der Geometrie* (*Fondamenti della geometria*) che dal 1899, anno in cui furono pubblicate, a oggi, hanno conosciuto oltre dieci edizioni e sono state tradotte praticamente in tutte le lingue. Già la data di pubblicazione da una parte pone questo volume simbolicamente come coronamento di un'epoca quanto mai feconda e decisiva per la ricerca matematica e per il problema dei fondamenti, e dall'altra ne fa quasi il punto di partenza ideale per ulteriori sviluppi sui quali i metodi e le concezioni in essa contenuti agiranno da potente catalizzatore.

L'antinomia di Burali Forti era già nota e Cantor aveva partecipato a Dedekind la scoperta del paradosso del massimo numero cardinale, ma sappiamo anche che Cantor stesso aveva avanzato una proposta di soluzione che nel complesso gli consentiva di non annettere importanza decisiva a tali antinomie. Frege, anche se misconosciuto, aveva già pubblicato il primo volume dei *Principi* e si accingeva a concludere la sua impresa. In sostanza, ad eccezione di particolari che nel complesso si potevano ritenere «trascura-

[47] David Hilbert nacque a Königsberg nel 1862. Compì i suoi studi nella città natale, salvo un semestre universitario passato all'università di Heidelberg. Si laurea nel 1884 con una dissertazione sulla teoria degli invarianti, argomento del quale continua a occuparsi sino al 1892; nel 1886 è libero docente, nel 1892 professore straordinario e nel 1894 ordinario sempre all'università di Königsberg. Nel 1895 viene chiamato all'università di Gottinga dove rimane per tutta la carriera accademica, pur ricevendo numerose altre chiamate da varie università, in seguito alla posizione di preminenza che l'estensione e la profondità della sua cultura matematica gli procurarono ben presto: Lipsia 1898, Berlino 1902, Heidelberg 1904. I suoi interessi possono essere classificati in pure ricerche matematiche che lo occupano sostanzialmente nel periodo dal 1892 al 1909 (anno della morte del suo migliore amico e collaboratore Hermann Minkowski, nato nel 1864) nell'ambito del quale tuttavia si innesta anche quell'aspetto delle ricerche sui fondamenti della matematica relativo ai fondamenti della geometria, circoscrivibile nel periodo 1893-1904; dopo il 1909 si interessa di fisica teorica, e dopo il 1916 di fondamenti della matematica. Muore a Königsberg nel 1943. Le opere di Hilbert che direttamente interesseranno la nostra esposizione verranno citate nel testo, in questo e nel prossimo capitolo.

bili», la riduzione della matematica all'aritmetica e quindi, almeno in senso lato, alla logica, poteva dirsi compiuta. Fuori da questo orizzonte (come ripetutamente abbiamo ricordato) rimaneva per Frege la geometria che era kantianamente costituita da giudizi sintetici a priori, il che è in qualche senso paradossale se si pensa che proprio dalla geometria aveva preso le mosse la consapevolezza dell'opportunità di distinguere fra applicabilità di una teoria e giustificazione della sua struttura logica, *genesi* di un concetto e sua *articolazione* matematica. È da questo punto di vista che le *Grundlagen* di Hilbert si pongono come completamento di un processo, nel senso che compiono una esplicita distinzione fra geometria in quanto *scienza dell'estensione*, per usare un'espressione di Padoa, e geometria come sistema ipotetico deduttivo inteso però in senso puramente formale, con le «ipotesi» iniziali del tutto indipendenti dal contenuto, empirico o no, che esse *eventualmente potranno* esprimere.

D'altra parte, dopo che la comunicazione di Russell a Frege aveva aperto «ufficialmente» la *crisi dei fondamenti* della matematica, il metodo assiomatico inteso in questo nuovo senso si presenta per la prima volta ancora con lo stesso Hilbert in *Über die Grundlagen der Logik und Mathematik* (*Sui fondamenti della logica e della matematica*) del 1904, come fecondo strumento per superare le difficoltà che sembravano intrinseche ai metodi fino ad allora escogitati per la fondazione della matematica; e come tale avrà applicazioni e affinamenti sorprendenti e decisivi nel nostro secolo, come vedremo nel prossimo capitolo.

Una volta che da parte matematica ne era stata segnalata l'esigenza, soprattutto grazie ai lavori di Dedekind e di Cantor, e quindi di Frege, il riconoscimento della natura «logica» degli assiomi dell'aritmetica si era nel complesso configurato come qualcosa di abbastanza «naturale»; un tale riconoscimento era in effetti molto più impegnativo e «innaturale» per quanto riguardava i principi della geometria, e trovava ben più gravi resistenze proprio per la innata convinzione che l'*oggetto* di quei principi fosse in fin dei conti pur sempre lo spazio fisico. Certo, le geometrie non euclidee avevano di molto contribuito a far vacillare la fede nell'assolutezza e evidenza dei principi della geometria ma non avevano portato in generale all'esigenza di una considerazione puramente logica di tali principi e delle loro relazioni. Si era indubbiamente evoluta la concezione delle proposizioni geometriche fonda-

mentali: accanto a coloro che ancora persistevano a considerarle come «evidenze», si parlava, con Riemann di «ipotesi» e con Helmholtz di «fatti» che stanno alla base della geometria; si intrecciavano fitte e impegnate discussioni sulla loro natura di proposizioni date kantianamente dalla pura intuizione, o frutto di una ipotetica postulazione (Riemann), o ottenute da una idealizzazione dell'esperienza (Helmholtz) o ancora come proposizioni trascendenti la realtà (Klein). Non era però cambiato in effetti quello che era ritenuto essere il riferimento fondamentale e comune di queste proposizioni: la geometria era pur sempre considerata come la scienza dello *spazio reale*, e la sua organizzazione deduttiva dipendeva strettamente da tale interpretazione privilegiata e *univoca*, sicché le sue proposizioni fondamentali, comunque le si considerasse, intuitive, ipotetiche o fattuali, avevano sempre, in definitiva, un «contenuto» precisamente individuato. Di nuovo, c'era l'idea della possibilità (esplicita in Riemann e Grassmann) di una teoria astratta delle forme di spazio di cui quello reale costituisce un esempio.

In quest'ottica, Dedekind e Cantor avevano mostrato come si potesse trattare aritmeticamente («logicamente») il continuo geometrico e anzi si erano anche espressi (si confronti la fine del paragrafo 2) per una concezione «astratta» del riferimento geometrico spaziale, ossia, nella sostanza, avevano chiaramente indicato la strada per una considerazione della geometria come «indipendente» dallo spazio fisico; ma ancora Frege, ad esempio, riteneva corretta la posizione di Kant circa le proposizioni geometriche come sintetiche *a priori* e individuanti relazioni e proprietà di ben precisi elementi spaziali in senso reale (i «punti», le «rette», ecc.). In forma diversa questo riferimento fondamentale allo spazio fisico è ravvisabile anche nelle concezioni di quegli stessi geometri che possono considerarsi come diretti precursori di Hilbert, e che anzi talora esprimono nei loro lavori metodi ed esigenze della nuova assiomatica in modo anche più chiaro ed esplicito dello stesso Hilbert: alludiamo in particolare a Moritz Pasch (1843-1930) e ad alcuni componenti di quella che può ben essere chiamata la «scuola italiana» ossia Giuseppe Peano, Giuseppe Veronese (1854-1917), Mario Pieri (1860-1913), Alessandro Padoa (1868-1937), Giovanni Vailati (1863-1909), Gino Fano (1871-1952), Federigo Enriques (1871-1946) e altri.

In tutti questi autori, in modi diversi si assiste ad una oscillazione tra una nuova concezione del rapporto fra realtà fisica e geometria e, più semplicemente a un diverso modo di concepire il ruolo del ragionamento formale nell'articolazione della conoscenza geometrica. In questo contesto era stato soprattutto lo sviluppo della geometria proiettiva, con il problema di distinguere proprietà metriche da proprietà proiettive, esplicitare il senso dell'introduzione di elementi ideali, fondare la validità di principi come quello di continuità di Poncelet o di dualità di Gergonne a portare ad un'analisi fine della natura assiomatica della geometria, ove «assiomatico» va inteso in un nuovo senso rispetto a quello euclideo. Entro questo panorama veniva a porsi la stessa questione delle geometrie non euclidee.

Felix Klein aveva tracciato nel 1872 il suo «programma di Erlangen» in cui erano pure considerazioni gruppali, ossia algebriche, che determinavano i rapporti delle varie geometrie; e Pasch, nel 1882, con le *Vorlesungen über die neuere Geometrie* (*Lezioni di geometria moderna*) aveva effettivamente dato una sistemazione «logica» in senso deduttivo alla geometria proiettiva, realizzando così il sogno accarezzato da von Staudt. L'eredità di Pasch venne raccolta nell'ultima decade del secolo dagli italiani: Peano nel 1889 traduceva in veste logico-simbolica le idee di Pasch ne *I principi della geometria logicamente esposti*; nel 1891 Veronese con i *Fondamenti di geometria a più dimensioni e a più specie di unità rettilinee* dava veste concreta (anche se nel complesso alquanto oscura e confusa) al primo sistema di geometria non archimedea e Pieri e Padoa, fra l'altro, ai congressi di matematica e di filosofia tenuti a Parigi nel 1900 esprimevano con chiarezza, rigore e organicità idee relative a una considerazione della geometria come sistema puramente logico. Anche se tutti questi autori ribadiscono la necessità di un collegamento tra pensiero geometrico e realtà fisica, enfatizzando la necessità dell'analisi della struttura logica della geometria, tendono a ribadire che questo collegamento non può essere univoco e deciso a priori.

L'innovazione delle *Grundlagen* hilbertiane sta nella *concreta* realizzazione di questa visione della geometria per cui non si dà per scontata una «realtà» che si suppone chiara ed evidente, oggetto di sensazioni, unica e incontrovertibile: le proposizioni fondamentali della geometria, gli assiomi che Hilbert pone alla base del suo sistema non

esprimono, in linea di principio, nessun contenuto che non sia quello dato dalle loro mutue relazioni di tipo puramente logico; la questione di come e se questi assiomi si applichino alla realtà diventa, come dice molto bene Freudenthal, «... esattamente la stessa che per ogni altra branca della matematica. Gli assiomi non sono verità evidenti. Essi non sono addirittura verità, nel senso usuale». Gli assiomi diventano «definizioni implicite» dei concetti e dei termini primitivi o indefiniti che essi contengono.

Malgrado una tradizione consolidata attribuisca ad Hilbert la paternità di una tale concezione, va detto che idee simili si ritrovavano esposte con estrema chiarezza già da Peano nel suo lavoro *I principi della geometria* sopra citati e ne *I fondamenti della geometria* del 1894, in cui, sempre ribadendo che la geometria ha a che fare con il mondo esterno, mostrava come diversi possano essere i concetti geometrici assunti come primitivi e come non sia né possibile né necessario in generale pretendere di definire questi ultimi. Assunti nel suo sistema come primitivi i concetti di «punto» e di «giacere tra *a* e *b*» ($c \varepsilon ab$), Peano afferma esplicitamente che «bisognerà determinare le proprietà dell'ente non definito *p* e della relazione $c \varepsilon ab$ mediante assiomi e postulati: l'osservazione la più elementare ci indica una lunga serie di proprietà di questi enti; a noi non resta che raccogliere queste cognizioni comuni, ordinarle, ed enunciare come postulati quelle sole che non si possono dedurre da altre più semplici». Sono gli assiomi quindi a determinare il significato delle nozioni primitive fornendo la base per dimostrare i teoremi veri su di esse. Le uniche definizioni possibili per Peano sono nominali e riguardano i concetti che possiamo definire *entro* una teoria a partire dai primitivi: «dire che l'oggetto, o nome *x*, si può definire, significa che combinando convenientemente le idee primitive coi termini di logica, si può formare un'espressione identica a quella indicata col nome *x*». Peano e la sua scuola furono particolarmente interessati alla questione concernente la scelta dei concetti primitivi e se Peano stesso mostrò per primo con alcuni esempi come fosse necessario, una volta che si considerino definizioni aggiuntive, garantire che queste definizioni non comportano la dimostrabilità di nuovi teoremi (è questa la condizione di non creatività delle definizioni) il suo allievo A. Padoa formulò un criterio generale per la definibilità di cui avremo occasione di parlare più avanti (capitolo V, paragrafo 2).

Malgrado il grande interesse di questi risultati è però innegabile che nella comunità matematica furono i *Fondamenti* di Hilbert a dettare le nuove linee di ricerca e la nuova concezione dell'assiomatica. Un rivolgimento di questo tipo non poteva certo non sollevare riserve e perplessità di ogni tipo, anche se, si noti bene, Hilbert non usa mai l'espressione «definizione implicita» che invece viene usata letteralmente dai geometri italiani prima nominati; egli in effetti tende a considerare i vari gruppi di assiomi che daremo più avanti come vere e proprie definizioni dei relativi concetti che in essi intervengono. Ma comunque si considerasse la cosa, che senso poteva avere «definire» dei concetti in modo non univoco? Come si poteva ad esempio pensare che venissero in tal modo definiti enti quali i punti, i piani, le rette, se di principio *qualunque* sistema di enti che soddisfacesse gli assiomi aveva lo stesso diritto di essere chiamato e inteso come costituito di punti, piani, rette? (e non è difficile nello stesso ambito geometrico trovare esempi di questo tipo).

Tipiche a questo proposito, addirittura irridenti a volte, le obiezioni di Frege in un nutrito scambio epistolare con Hilbert, e assai illuminanti, d'altra parte, le risposte che Hilbert dà a queste obiezioni. Frege sostanzialmente rimprovera a Hilbert di confondere i significati di termini quali «assioma», «definizione», «spiegazione», e in definitiva di non giungere, con i suoi assiomi, a «definire» alcunché per il semplice motivo che egli, in realtà, assume come noti i significati dei termini «punto», «retta», «piano», ecc. che figurano negli assiomi. A questo Hilbert risponde di non volere «presupporre nulla come noto» e di vedere negli assiomi «la definizione dei concetti punto, retta, piano... Se si cercano altre definizioni di "punto", ricorrendo ad esempio a perifrasi come "privo di estensione" ecc.,» prosegue Hilbert, «si capisce che debbo oppormi nel modo più deciso a siffatti tentativi; si va infatti alla ricerca di qualcosa là dove non la si potrà mai trovare, per il semplice motivo che non è dove la si cerca; e così facendo tutto si disperde, diviene vago e confuso e degenera in un giocare a rimpiattino».

Altro punto cruciale, legato al precedente, ma che mette in evidenza il significato nuovo che certe proprietà dei sistemi assiomatici venivano acquistando nella concezione hilbertiana è quello circa la natura logica degli assiomi e l'esistenza di «modelli» degli

assiomi stessi. Per Frege gli assiomi (in particolare, della geometria) sono proposizioni «... vere; ma che non vengono dimostrate perché la loro conoscenza scaturisce da una fonte conoscitiva di natura extralogica, che possiamo chiamare intuizione spaziale. Il fatto che gli assiomi siano veri ci assicura di per sé che essi non si contraddicono fra loro, e ciò non abbisogna di alcuna ulteriore dimostrazione...». Proprio in risposta a questa frase Hilbert ribatte: «Mi ha molto interessato leggere... questa frase, perché io, da quando ho cominciato a riflettere, scrivere e tenere conferenze su questo argomento, ho detto esattamente il contrario: se assiomi arbitrariamente stabiliti non sono in contraddizione, con tutte le loro conseguenze, allora essi sono veri, allora esistono gli enti definiti per mezzo di quegli assiomi. Questo è per me il criterio della verità e dell'evidenza.»

In risposta a ulteriori obiezioni di Frege, Hilbert rileva: «Lei dice che i miei concetti, per esempio "punto", e "fra", non sono stabiliti univocamente... Certamente si comprende da sé che ogni teoria è solo un telaio, uno schema di concetti unitamente alle loro mutue relazioni necessarie e che gli elementi fondamentali possono venir pensati in modo arbitrario. Se... voglio intendere un qualunque sistema di enti, per esempio il sistema: amore, legge, spazzacamino, ..., allora basterà che assuma tutti i miei assiomi come relazioni fra questi enti perché le mie proposizioni, ad esempio il teorema di Pitagora, valgano anche per essi. In altre parole: *ogni teoria può essere sempre applicata a infiniti sistemi di enti fondamentali...* La circostanza ora menzionata... non può mai rappresentare un difetto di una teoria (ne è piuttosto un grandissimo pregio) e in ogni caso è inevitabile» [corsivo nostro].

Nelle parole precedenti è sostanzialmente contenuto il senso del «formalismo» hilbertiano (e viene naturale pensare ad alcune analoghe affermazioni di Dedekind), che verrà tuttavia affinandosi e precisandosi ulteriormente negli anni successivi e, come vedremo nel prossimo capitolo, troverà nuova espressione attorno agli anni venti. E Hilbert in effetti dovrà superare ostacoli non indifferenti per imporre al mondo matematico queste sue vedute; in compenso esse acquisirono una portata del tutto generale e influenzarono enormemente il modo stesso di intendere la «matematizzazione» della realtà. Sono significative a questo proposito le parole che Einstein ebbe a scrivere nel 1921, nel suo saggio *Geo-*

metria ed esperienza: «Nella misura in cui i teoremi matematici si riferiscono alla realtà essi non sono sicuri, e nella misura in cui sono sicuri essi non si riferiscono alla realtà... Il progresso portato dall'assiomatica consiste nella decisa separazione della forma logica e del contenuto intuitivo e reale... Gli assiomi sono creazioni volontarie della mente umana... Io assegno grande importanza a questa concezione della geometria perché se non avessi preso familiarità con essa, non sarei mai stato in grado di sviluppare la teoria della relatività».

La parola d'ordine che emerge non solo nella logica o nella matematica del tempo è che non sono gli oggetti specifici (figure, corpi, insiemi, ecc.) che contano, ma le loro relazioni, le *strutture* che essi formano, che danno significato e possibilità di sviluppo concettuale alla matematica. La realtà dei punti, delle figure, ecc. non può essere direttamente oggetto di pensiero esatto. È la svolta *strutturalista* di cui Dedekind era stato uno dei primi proponenti e di cui l'assiomatica di Hilbert costituisce il versante per così dire logico-linguistico. Nel giro di pochi decenni non solo la teoria degli insiemi per opera di Zermelo, ma le varie teorie algebriche, dei gruppi, dei corpi, dei campi, ecc., la teoria della misura, l'Analisi, la topologia, tutta intera la matematica insomma, avrebbero assunto veste assiomatica.

Ma ritorniamo alla geometria nella formulazione hilbertiana. Hilbert formula gli assiomi della geometria tridimensionale in termini di relazioni che riguardano tre tipi diversi di oggetti: i punti, le rette e i piani che indica rispettivamente con A, B, C, ..., a, b, c, ..., α, β, γ, ... Oltre a questi tre concetti primitivi, Hilbert assume la relazione di *collegamento* o *incidenza* che sussiste fra un punto A e una retta a quando A giace su a o a passa per A (e analogamente nel caso di punti e piani o rette e piani), la relazione di «stare fra» la cui prima introduzione risale a Pasch e che sussiste fra tre punti A, B, C, che giacciono su una stessa retta, quando B si trova fra A e C, la relazione di *congruenza* tra segmenti e che viene successivamente estesa ad angoli, dove tanto il concetto di segmento che quello di angolo vengono definiti a partire dai primitivi. Il ruolo degli assiomi è di fissare il «significato» di questi concetti primitivi determinandone le relazioni fondamentali. Hilbert distingue cinque gruppi di assiomi ed il raggruppamento ha il senso di isolare i diversi tipi di situazioni che si possono presentare nello spazio:

fatti sull'incidenza, sulla congruenza, sullo stare fra, ecc. Lo scopo dell'assiomatizzazione è proprio quello di mostrare i possibili rapporti fra questi fatti, le alternative e le dipendenze. Diamo ora gli assiomi di Hilbert con l'avvertenza che non si tratta di quelli della prima edizione del 1899 ma di quelli tratti dalla decima edizione delle *Grundlagen*.

Gruppo I – Sono 8 assiomi: 3 piani e 5 spaziali. Essi sono detti di *collegamento*.

Assiomi piani:
I 1. Per due punti A e B, c'è sempre una retta che appartiene ad ognuno dei due punti A e B.
I 2. Per due punti A e B c'è al massimo una retta che appartiene ad ognuno dei due punti A e B.
I 3. Su una retta ci sono sempre almeno due punti. Ci sono almeno tre punti che non giacciono su una retta.

Assiomi spaziali:
I 4. Per tre punti qualsiasi A, B, C che non giacciono su una stessa retta, c'è sempre un piano α che appartiene ad ognuno dei tre punti A, B, C. Per ogni piano c'è sempre un punto che gli appartiene.
I 5. Per tre punti qualsiasi A, B, C che non giacciono su una medesima retta c'è al massimo un piano che appartiene a ciascuno dei tre punti A, B, C.
I 6. Se due punti A, B di una retta a giacciono in un piano α, allora ogni punto di a è nel piano α.
I 7. Se due piani α, β hanno in comune un punto A, allora hanno in comune almeno un altro punto B.
I 8. Ci sono almeno quattro punti che non stanno su un piano.

Gruppo II – Sono in tutto 4 assiomi detti di *ordinamento*:
II 1. Se un punto B giace fra un punto A ed un punto C, allora A, B, C sono tre punti distinti di una retta e B giace pure fra C ed A.
II 2. Per ogni due punti A e C, c'è sempre almeno un punto B, sulla retta AC, tale che giace fra A e B.

II 3. Di tre punti qualsiasi di una retta ce n'è al massimo uno che giace fra gli altri due.

II 4. Siano A, B, C tre punti non allineati ed a una retta del piano ABC che non passi per alcuno dei punti A, B, C: allora, se la retta a passa per un punto del segmento AB, essa passa certamente anche per un punto del segmento AC, ovvero per un punto del segmento BC.

Gruppo III – Sono in tutto 5 assiomi detti di *congruenza*:

III 1. Se A, B sono due punti di una retta a ed inoltre A' è un punto sulla stessa retta ovvero su un'altra a', si può sempre trovare un punto B', da una data parte della retta a' rispetto ad A', tale che il segmento AB sia congruente, ovvero uguale, al segmento $A'B'$, in simboli:

$$AB \equiv A'B'$$

III 2. Se un segmento $A'B'$ ed un segmento $A''B''$ sono congruenti ad uno stesso segmento AB, allora anche il segmento $A'B'$ è congruente al segmento $A''B''$, ovvero brevemente: se due segmenti sono congruenti ad un terzo, essi sono congruenti fra loro.

III 3. Siano AB e BC due segmenti senza punti in comune su una retta a ed $A'B'$ e $B'C'$ due segmenti sulla medesima retta o su un'altra retta a', sempre senza punti in comune; allora se è:

$$AB \equiv A'B' \text{ e } BC \equiv B'C'$$

è pure:

$$AC \equiv A'C'$$

III 4. Siano dati un angolo $\sphericalangle\ (h, k)$ in un piano α ed una retta a' in un piano α' come pure un determinato lato di a' in α'. Si indichi con h' una semiretta della retta a', che abbia origine nel punto O': c'è allora nel piano α' una e una sola semiretta k', tale che l'angolo $\sphericalangle\ (h, k)$ è congruente ovvero uguale all'angolo $\sphericalangle\ (h', k')$ ed allo stesso tempo tutti i punti interni all'angolo $\sphericalangle\ (h', k')$ stanno dalla parte di a', in simboli:

$$\sphericalangle\ (h, k) \equiv \sphericalangle\ (h', k').$$

Ogni angolo è congruente a se stesso, cioè è sempre:

$$\sphericalangle\ (h, k) \equiv \sphericalangle\ (h, k).$$

III 5. Se per due triangoli ABC e $A'B'C'$ valgono le congruenze:

$$AB \equiv A'B',$$
$$AC \equiv A'C',$$
$$\sphericalangle\ BAC \equiv \sphericalangle\ B'A'C'$$

allora è sempre valida anche la congruenza:

$$\sphericalangle ABC \equiv \sphericalangle A'B'C'.$$

Gruppo IV – Assioma delle *parallele*:

IV. Siano *a* una qualsiasi retta ed *A* un punto fuori di *a*: allora c'è, nel piano definito da *A* e da *a*, al massimo una retta che passa per *A* e che non interseca la *a*.

Gruppo V – Sono due assiomi detti di *continuità*:

V 1. (*Assioma della misura, ovvero assioma archimedeo.*) Se *AB* e *CD* sono due segmenti qualsiasi, c'è un numero *n* tale che il trasporto del segmento *CD* reiterato *n* volte da *A* sulla semiretta passante per *B*, porta al di là del punto *B*.

V 2. (*Assioma di completezza lineare.*) Il sistema dei punti di una retta con le sue relazioni di ordinamento e congruenza non è suscettibile di un ampliamento per il quale rimangano inalterate le relazioni sussistenti tra gli elementi precedenti come pure le proprietà fondamentali di ordinamento lineare e congruenza che seguono dagli assiomi I-III ed anche V 1.

Sono forse opportune alcune osservazioni sul ruolo degli assiomi e sulla differenza con sistemazioni posteriori forse più familiari al lettore. Hilbert assume come fondamentali i concetti di «punto», «retta» e «piano» ed è costretto dunque ad introdurre una relazione come quella di incidenza; in molte delle sistemazioni assiomatiche posteriori ciò viene evitato assumendo come primitivo il solo concetto di punto e definendo rette e piani come particolari insiemi di punti (sostituendo così all'incidenza l'ordinaria relazione insiemistica di appartenenza). Per quanto riguarda la relazione «stare fra» essa costituisce un esempio tipico di come una relazione fondamentale fra enti geometrici possa passare del tutto inosservata all'intuizione; il suo ruolo centrale sta nel fatto che ci permette di introdurre ordinamenti su una retta senza assumerne uno a priori ed è in base ad essa ancora che possiamo definire i concetti di intervallo, di semiretta, ecc., e in generale distinguere fra interno ed esterno di una figura chiusa. Prima di Pasch tutti questi fatti venivano usati inconsapevolmente. L'assioma II.4 in particolare, noto come *assioma di Pasch*, stabilisce una proprietà centrale riguardo ai segmenti che intersecano un lato di un triangolo (determinato dai tre punti *A*, *B*, *C*): se una retta *a* entra in un

triangolo senza passare per nessuno dei vertici ed interseca il lato AB essa dovrà intersecare anche uno degli altri due lati.

Passando agli assiomi di congruenza, il fatto da porre in luce è che assumendo questa relazione come fondamentale e come base per costruire tutta la dimensione metrica della geometria, Hilbert riesce da una parte a eliminare ogni ricorso alla sovrapponibilità (e quindi al movimento) come faceva Euclide e dall'altra ad evitare la via rigorosa per introdurre questa nozione, attraverso la teoria dei gruppi. Negli assiomi si parla di congruenza non solo tra segmenti ma anche fra angoli: si tratta di concetti che possono essere definiti, il primo in termini della relazione «stare fra», il secondo come coppia di semirette aventi un'origine in comune. L'assioma IV delle parallele non è che una delle formulazioni più note del quinto postulato euclideo.

Ben maggiore interesse hanno gli assiomi del quinto gruppo che introducono uno dei temi centrali della ricerca che Hilbert conduce nelle *Grundlagen*. Si tratta degli assiomi di continuità di cui il primo, il postulato di Archimede, Hilbert chiamava più propriamente postulato della misura in quanto ci garantisce la possibilità, dati due segmenti AB e CD, di «misurare» l'uno prendendo come unità di misura l'altro. Nel corso dell'opera Hilbert mostra come appunto l'assioma di Archimede possa essere realizzato costruendo una geometria analitica su un campo numerico che non coincide con quello dei reali, consiste solo di numeri algebrici e non contiene le radici quadrate di ogni numero. Per formulare la continuità nel senso in cui oggi la si intende, occorre perciò un ulteriore assioma che in termini di retta numerica stabilisce la sezionabilità in senso dedekindiano. Il problema di trovare una formulazione geometrica per la continuità è legato così a quello di trovare una formulazione che caratterizzi il sistema dei reali: Hilbert affronta simultaneamente i due problemi nel lavoro *Über den Zahlbegriff* (*Sul concetto di numero*, 1900) in cui si occupa appunto dei numeri reali e nella seconda edizione delle *Grundlagen* in cui introduce il secondo assioma di continuità, l'assioma di completezza, in una forma più forte di quella in cui l'abbiamo dato (e la cui equivalenza con tale forma, una volta assunti anche gli altri assiomi, è stata dimostrata da P. Bernays).

L'assioma stabilisce una proprietà di massimalità per i modelli degli altri assiomi e questa massimalità è formulata in termini

«metamatematici» in quanto afferma che modello dell'assioma saranno tutti e soli i modelli degli altri assiomi – che indicheremo con H – che non hanno estensioni proprie che siano ancora modelli di H. In termini intuitivi: la completezza afferma che lo spazio ha il massimo numero di punti, rette, piani, possibili compatibile col verificare gli assiomi.

Con un analogo assioma, nell'articolo del 1900 sopra citato Hilbert formulerà la completezza del sistema dei numeri reali. La cosa è importante da due punti di vista almeno. In primo luogo questo modo di formulare la continuità geometrica (Archimede più completezza) permette di superare una difficoltà che già von Staudt aveva incontrato cercando di dimostrare il teorema fondamentale della geometria proiettiva secondo il quale per determinare una corrispondenza proiettiva è sufficiente determinarla per tre punti; questo teorema serviva a von Staudt per edificare l'intera geometria proiettiva senza far ricorso alla nozione di misura, ma a sua volta, per essere dimostrato, necessitava di un principio di continuità. Come allora determinare un principio di continuità senza ricorrere a rappresentazioni numeriche? La risposta di Hilbert, con l'assioma di completezza, porta in primo piano l'intergioco fra teorie e modelli. Questo ci porta al secondo aspetto di interesse dell'assioma. Lo scopo degli assiomi di Hilbert è quello di *caratterizzare* la struttura dello spazio euclideo. Questo significa, se prendiamo la richiesta in senso forte, richiedere che la teoria sia *categorica* (nel senso, introdotto da J. Dewey ma per la prima volta usato in contesto matematico da O. Veblen in un suo lavoro del 1905) vale a dire che tutti i suoi modelli siano isomorfi fra di loro, abbiano cioè la stessa struttura.

Ora, Hilbert riesce, attraverso l'introduzione di un calcolo dei segmenti, a definire in termini puramente geometrici un sistema di grandezze che possono servire come coordinate per lo spazio all'interno di ogni modello e a provare che questo sistema è isomorfo a quello dei numeri reali e che quindi ogni modello si ottiene prendendo come punti le terne di numeri reali, come rette e piani i luoghi di punti definiti nel modo solito nella geometria analitica. In questa prospettiva il ruolo dell'assioma di completezza è quello di garantire la categoricità, ed esso fornisce così il primo esempio di quegli *assiomi estremali*, come verranno chiamati da Carnap e F. Bachmann in un loro articolo del 1936, che permettono di isolare

e spesso di caratterizzare a meno di isomorfismi strutture massimali e minimali. Tale ad esempio è l'assioma di induzione, che stabilisce una condizione di minimalità per la struttura dei naturali; e tale cercherà di essere l'assioma di fondazione nella teoria degli insiemi per Fraenkel. Lungo questa linea si giungerà ad esaminare queste proprietà di massimalità e minimalità in contesti più ampi, e al concetto di model-completezza introdotto da A. Robinson negli anni cinquanta.

È importante comunque notare che se l'assioma di completezza ha una chiara evidenza intuitiva, esso presenta una forma anomala rispetto agli altri assiomi per la già detta forma metamatematica, cosicché già nel 1928 fu sottoposto a critiche da parte di R. Baldus e fu proprio da queste critiche che partirono le ricerche di Carnap e Bachmann sulle proprietà logiche degli assiomi estremali. Va comunque aggiunto che per quanto riguarda la geometria è possibile di fatto rimpiazzare l'assioma di completezza con un altro, l'assioma di continuità di Cantor, che si riferisce non ai modelli degli assiomi, ma direttamente agli oggetti geometrici.

In modo del tutto naturale siamo venuti a parlare delle proprietà «metamatematiche» del sistema di assiomi di Hilbert. È questo aspetto che per Hilbert mostra l'utilità del metodo assiomatico: egli ribadisce più volte che gli assiomi fissano i dati della nostra intuizione e i diversi gruppi d'assiomi rappresentano aspetti diversi di questa intuizione. In questa prospettiva lo studio delle compatibilità e delle dipendenze tra gli assiomi costituisce – come afferma Hilbert – una «analisi logica della nostra intuizione spaziale». In altre parole, l'esigenza che muoveva Hilbert non era un'esigenza di puro rigore, ma il tentativo di analizzare, con nuovi metodi, i *fatti* alla base della geometria. Così, per limitarci ad alcuni esempi, Hilbert indagò sistematicamente il ruolo dell'assioma di Archimede, come pure la possibilità di sviluppare la geometria proiettiva in termini autonomi utilizzando i teoremi di Pascal e di Desargues. Il metodo era sempre lo stesso: costruire modelli che mostravano come una data proposizione potesse essere vera senza che ne fosse vera un'altra (che risultava così indipendente dalla prima). Così, ad esempio, il teorema di Pascal si può dimostrare senza l'assioma di Archimede ma con l'aiuto degli assiomi di congruenza (che sono di carattere metrico) ma non è ottenibile senza almeno uno dei due tipi d'assiomi; in modo analogo, il teo-

rema di Desargues per il piano non è ottenibile senza usare gli assiomi di congruenza, a meno di non assumere un assioma di tridimensionalità. Furono proprio esempi come questi a rivelare al mondo matematico come il metodo assiomatico non fosse una pura questione d'eleganza ma avesse una rilevanza matematica centrale. Quale fosse lo stupore lo può indicare la recensione di Poincaré alle *Grundlagen*, che pure veniva da un matematico di professione convenzionalista e che tuttavia ad ogni passo confessa la sua meraviglia di fronte alla pluralità di geometrie possibili che il lavoro di Hilbert mostrava.

Non è il caso in questa sede di approfondire il discorso sugli sviluppi che l'indagine sui fondamenti della geometria ebbe dopo il lavoro di Hilbert; ci basti ricordare O. Veblen che nel 1905 tentò un'assiomatizzazione, basata sul solo concetto di «stare fra», tanto della geometria euclidea che di quella iperbolica, o E.V. Huntington che utilizzò invece come nozioni primitive, nel 1913, i concetti di sfera e di inclusione, o infine le ricerche della scuola italiana, in particolare di M. Pieri che si occupò specificamente dei fondamenti della geometria proiettiva.

Come già abbiamo accennato, il valore delle *Grundlagen* hilbertiane non si limita alla geometria, ma riguarda il senso stesso della nuova accezione di assiomatica che vi veniva prospettata. Due sono i ruoli che il metodo può avere e avrebbe avuto. Uno è quello di permettere la formulazione di teorie per cui non si ha né si intende avere un modello privilegiato: questo vale per gruppi, anelli, corpi, ecc. in genere per tutte le strutture astratte; per oggetti di questo tipo il metodo assiomatico è l'unico adeguato in quanto ci permette di prescindere dalle possibili realizzazioni concrete per concentrarci sulla struttura astratta. L'altro è un ruolo fondazionale: qui si tratta di caratterizzare strutture specifiche quali i naturali, i reali, lo spazio euclideo, ecc.

In entrambi i casi diventa essenziale un'indagine che ha per oggetto gli stessi sistemi assiomatici, vale a dire un'indagine *metamatematica* che ne saggi l'adeguatezza ai fini proposti. La prima questione è quella sull'esistenza di modelli per un sistema assiomatico che è strettamente legata a quello della *coerenza* o *noncontraddittorietà* dello stesso, vale a dire la possibilità di ottenere per via logica, dai suoi assiomi, proposizioni contraddittorie, del tipo $\mathscr{A} \wedge \neg \mathscr{A}$. Il legame è dato dal fatto immediato che un enunciato $\mathscr{A}$ deriva da

una teoria $\mathfrak{T}$ se e solo se non c'è modello di $\mathfrak{T}$ che non sia anche modello di $\mathscr{A}$ cosicché per provare che $\mathscr{A}$ *non* deriva da $\mathfrak{T}$ basta trovare un modello di $\mathfrak{T}$ che non sia modello di $\mathscr{A}$ cioè un modello che soddisfi $\mathfrak{T}$ e $\neg \mathscr{A}$. Questo mostra quanto strettamente connessi ai problemi di coerenza siano anche i problemi di *indipendenza*: dire che $\mathscr{A}$ è indipendente da $\mathfrak{T}$ significa che né $\mathscr{A}$ né $\neg \mathscr{A}$ scendono da $\mathfrak{T}$ e questo comporta che l'indipendenza di $\mathscr{A}$ coincide con la soddisfacibilità tanto di $\mathfrak{T}$ e $\mathscr{A}$ quanto di $\mathfrak{T}$ e $\neg \mathscr{A}$.

Tutte queste indagini utilizzano essenzialmente la costruzione di interpretazioni che permettono di riportare lo studio dei modelli delle teorie in esame o a situazioni «concrete» (questo soprattutto nel caso di interpretazioni *finite*) o di tipo analitico (numeri complessi, algebrici, reali, ecc.). In questo senso tutte le dimostrazioni di coerenza che vengono date sono *relative*, non *assolute*, ossia rimandano alla coerenza di altre teorie accettate come tali (salvo ovviamente il caso delle interpretazioni finite di natura concreta). Non troviamo in nessuno di questi autori né l'idea che sia possibile un'analisi delle dimostrazioni in sé e per sé, né la determinazione univoca di un ambito, un universo in cui tutti i possibili modelli siano reperiti. Questo spiega spesso le ambiguità e le confusioni che si possono trovare riguardo concetti basilari come completezza, categoricità ed indipendenza. Non esiste in altre parole quella rigida distinzione tra aspetto sintattico e semantico che per noi oggi è abituale.

Ciò non toglie che un'esigenza di un'analisi *diretta* delle proprietà metamatematiche delle teorie su uno sfondo unitario si presenti in diverse forme. Al congresso di Parigi del 1900, Hilbert porrà esplicitamente in testa alla sua lista di 23 problemi due problemi di carattere tipicamente metamatematico: il problema dell'ipotesi del continuo di Cantor – di cui abbiamo già parlato a suo tempo (paragrafo 3) – e quello di provare la coerenza degli assiomi dell'aritmetica (che Hilbert intende come teoria dei numeri reali). Quest'ultimo problema è importante perché era proprio a questa teoria che essenzialmente rimandavano tutte le interpretazioni che Hilbert aveva dato per i vari sistemi geometrici considerati. Si trattava quindi, in questo caso, di trovare una dimostrazione che non rimandasse ad altre teorie, e quindi una dimostrazione in qualche senso «assoluta». Hilbert afferma esplicitamente che «la dimostrazione della noncontraddittorietà degli assiomi è anche la

dimostrazione dell'esistenza matematica dell'aggregato dei numeri reali, ovvero del continuo», ribadendo quel collegamento coerenza-esistenza, che Frege (e poi molti altri) non accettava come ovvio.

Chi, partendo invece da Frege, non accetta i termini hilbertiani di questa equazione saranno d'altra parte Russell e Whitehead che nei *Principia mathematica* codificheranno quella Grande Logica come teoria delle proposizioni e dei concetti in cui *tutta* la matematica (geometria inclusa) dovrà essere ricostruibile. Per loro il dominio è assegnato e univoco e non si tratta né di analizzare le dimostrazioni quando si indagano problemi di coerenza, né di interpretare in altre teorie, ma solo di ricavare dai principi della Grande Logica un enunciato che afferma l'esistenza di interpretazioni all'interno dell'universo che soddisfano specifiche funzioni proposizionali che rappresentano gli assiomi della teoria matematica in esame. Da questo punto di vista i *Principia* costituiscono l'antitesi delle *Grundlagen* nel senso che negano *a priori* la possibilità stessa di indagini metamatematiche sulla coerenza, categoricità, indipendenza degli assiomi assunti.

Paradossalmente, in questo tentativo di ricondurre a pochi principi fondamentali tutta intera la matematica ha un ruolo fondamentale l'intervento di G. Peano che nel suo progetto del *Formulario* uscito in diverse edizioni sempre ampliate dal 1898 al 1908, si era posto esplicitamente l'obiettivo di individuare l'ideografia necessaria per analizzare e trascrivere concettualmente tutta intera la matematica pur in una prospettiva certo differente da quella di Frege. Il riferimento esplicito di Peano è a Leibniz e alla sua *lingua speciosa* più che all'ideografia di Frege – che pure Peano conobbe e che giudicò traducibile nella propria. Per Peano non si tratta di *fondare* l'aritmetica né qualunque teoria specifica nel senso di Frege o di Hilbert, ma di analizzare in tutti i possibili teoremi matematici i concetti fondamentali, isolarli, studiarne i rapporti di definibilità, ed infine, individuati i primitivi, simbolizzarli in modo univoco e redigere un vero e proprio catalogo di tutti i teoremi matematici interessanti.

In questo programma la logica matematica – è proprio questo il termine che usa Peano – interviene fornendo nozioni logiche fondamentali (classe, appartenenza, identità, implicazione formale, variabile, connettivi e quantificatori ecc.) mediante i quali è possi-

bile questa trascrizione unitaria. Peano è perfettamente consapevole che la logica matematica non è solo questo e che il *Formulario* rappresenta il livello elementare dell'indagine logica, ma il suo obiettivo è eminentemente pratico, quello di fornire una sorta di Enciclopedia, che unifichi in qualche modo un sapere matematico che sempre più si andava sviluppando e diversificando (e quindi spesso con varianti ambigue nelle definizioni o nello stesso simbolismo). In altre parole il *Formulario* non è un tentativo di fondazione unitaria come i *Principia* ma è innegabile – come Russell non perse mai occasione di ribadire – che la precisione delle analisi concettuali di Peano fu un aiuto inestimabile nella compilazione degli stessi *Principia*, dal momento che mostrava in concreto, per larghe parti della matematica, come poche nozioni logiche fossero in grado di analizzare la massa eterogenea ed enorme dei contenuti matematici specifici: «Sino a che non incontrai Peano», ebbe a scrivere Russell in una lettera a Jourdain, «non mi passò mai per la testa che la logica simbolica sarebbe potuta essere di qualche uso per i principi della matematica, poiché conoscevo l'armamentario booleano e lo trovavo privo di utilità».

Di fatto però, dopo quest'incontro con Peano, fu a Frege che Russell si rivolse alla ricerca di una *fondazione* per la matematica, che non fosse una semplice unificazione linguistico-concettuale, come avviene in certo senso per la maggior parte dei contributi di tutta la scuola di Peano, salvo quelli più specificamente riguardanti i fondamenti della geometria: malgrado la ricchissima messe di analisi puntuali e in molti casi rivelatrici, quello che ad essi mancava era una prospettiva generale di più ampio respiro che accettasse i legami con le problematiche filosofiche.

Un'altra eredità delle *Grundlangen* di Hilbert era il problema di determinare cosa significa *caratterizzare* una struttura mediante un sistema assiomatico. Un modo che abbiamo visto è quello forte che identifica caratterizzabilità con categoricità. Come però già poco dopo la pubblicazione delle *Grundlagen* fu sottolineato da più studiosi, a questa condizione se ne poteva sostituire una più debole che non mirava alla unicità della struttura modello, quanto alla «decidibilità» di ogni enunciato del linguaggio in base agli assiomi. La proprietà in questione è la *completezza* (*sintattica*) di cui una teoria gode quando per ogni enunciato $\mathscr{A}$ del linguaggio o $\mathscr{A}$ o la sua negazione (ma non ovviamente entrambi) scendono dagli as-

siomi. È chiaro che sotto ragionevoli ipotesi la categoricità implica la completezza ma non è ovvio, e anzi una volta precisati i termini si può stabilire che non vale, il viceversa. Nella letteratura dedicata all'assiomatica fino sostanzialmente agli anni trenta c'è spesso un'ambiguità fra questi due concetti che dipende da una non chiara visione dei rapporti fra aspetto semantico e aspetto sintattico e dalla non esplicitazione del tipo di linguaggio usato. Ciò non toglie che a partire dalle *Grundlagen* ci sia stato un vero e proprio fiorire di indagini metamatematiche oggi dimenticate ma estremamente interessanti sia per i risultati specifici che per l'esplorazione di proprietà metamatematiche nuove legate alla categoricità e all'indipendenza come «l'indipendenza completa» (indagata da A. Church) la «primalità» (introdotta da H.M. Sheffer), altre forme di completezza (studiate da E. Moore e H.C. Langford), ecc. Nei primi anni del secolo si distinsero in queste indagini e in quelle parallele riguardanti la definibilità e la scelta dei concetti primitivi, soprattutto la scuola di Harvard e la scuola italiana, come abbiamo visto nel capitolo II parlando dell'algebra della logica.

E proprio a questo proposito vogliamo fare un'ultima osservazione, di natura per così dire «esterna». Nei due congressi del 1900 sopra ricordati, la scuola metamatematica italiana, capeggiata da Peano, aveva indubbiamente la preminenza in campo europeo. Per Russell – abbiamo visto – l'incontro con Peano e con i suoi allievi significò addirittura la rivelazione della possibilità di un approfondimento logico dell'analisi che lo determinò nell'indirizzare in questo senso le sue ricerche. Orbene, nel prossimo capitolo dovremo tentare di spiegarci come sia potuto avvenire quanto denuncia Hans Freudenthal e cioè che «il trionfo degli studiosi italiani dei fondamenti ricorda la vittoria di Pirro. Dopo le *Grundlagen der Geometrie* essi cessarono di interessarsi di fondamenti della geometria, e dopo i *Principia mathematica* essi diedero l'addio alla logistica».

BIBLIOGRAFIA

F. Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea, New York 1967 [edizione originale 1926-27].

U. Bottazzini, *Il calcolo sublime*, Boringhieri, Torino 1981.

F. Waismann, *Introduzione al pensiero matematico*, Boringhieri, Torino 1971 [edizione originale 1936].

AA.VV., *From Frege to Gödel. A source book in mathematical logic 1879-1931*, a cura di Jean van Heijenoort, Harvard University Press, Cambridge 1967.

G. Cantor, *Abhandlungen mathematischen und philosophischen Inhalts*, Olms Verlag, Hildesheim 1966.

Id., *Contributions to the founding of the theory of transfinite numbers*, Dover Publications, Inc., New York 1915.

J. Cavaillès, *Remarques sur la formation de la théorie abstraite des ensembles*, Hermann, Parigi 1938.

P.R. Halmos, *Teoria elementare degli insiemi*, Feltrinelli, Milano 1970 [edizione originale 1960].

A. Abian, *La teoria degli insiemi e l'aritmetica transfinita*, Feltrinelli, Milano 1972 [edizione originale 1965].

K. Kuratowski e A. Mostowski, *Set theory*, North Holland, Amsterdam 1976^2.

J.W.R. Dedekind, *Scritti sui fondamenti della matematica*, a cura di Francesco Gana, Bibliopolis, Napoli 1982.

P. Dugac, *Richard Dedekind et les fondements des Mathématiques*, Vrin, Parigi 1976.

G. Frege, *Logica e aritmetica*, a cura di Corrado Mangione, Boringhieri, Torino 1965.

Id., *Scritti Postumi*, a cura di Eva Picardi, Bibliopolis, Napoli 1986.

Id., *Alle origini della logica*, Epistolario, a cura di Corrado Mangione, Boringhieri, Torino 1983.

Id., *Ricerche logiche*, a cura di Michele De Francesco, Guerini e associati, Milano 1988.

D. Hilbert, *I fondamenti della geometria*, con i supplementi di P. Bernays, Feltrinelli, Milano 1970.

Id., *Ricerche sui fondamenti della matematica*, a cura di Michele Abrusci, Bibliopolis, Napoli 1978.

J. Cavaillès, *Méthode axiomatique et Formalisme*, Hermann, Parigi 1938.

C. Reid, *Hilbert*, Springer Verlag, Berlino 1978.

H. Freudenthal, «The main trends in the foundations of geometry in the 19th century», in *Logic methodology and philosophy of science*, Stanford University Press, Stanford 1962.

G. Peano, *Opere scelte*, a cura di Ugo Cassina, Edizioni Cremonese, Roma, 3 voll., 1957-1959.

Id., *Formulario Mathematico*, (a cura di Ugo Cassina), Edizioni Cremonese, Roma 1960.

H. Kennedy, *Peano. Storia di un matematico*, Boringhieri, Torino 1983 [edizione originale 1980].

H. Poincaré, *Sui fondamenti della geometria*, Editrice la Scuola, Brescia 1990 [edizione originale 1898].

Id., *Opere epistemologiche*, Piovan editore, Abano Terme 1989, vol. I.

B. Russell, *An essay on the foundations of Geometry*, Dover Publisher Inc., New York 1956 [edizione originale 1897].

CAPITOLO QUARTO
LE ANTINOMIE E IL PROBLEMA DEI FONDAMENTI

1. LA CRISI DEI FONDAMENTI

Con la pubblicazione delle *Grundlagen* di Hilbert e dei *Grundgesetze* di Frege, la grande stagione della ricerca logica del XIX secolo può dirsi conclusa: punto d'arrivo di tutto un secolo di indagini sui fondamenti della matematica, queste opere, malgrado il valore paradigmatico che non tardarono ad assumere, non rappresentarono però il punto di partenza diretto della ricerca logica del nuovo secolo. Tra di esse e i lavori di Russell e Whitehead, il «secondo» Hilbert e Brouwer, si frappone un diaframma che segna una frattura e un deciso mutamento dell'orizzonte problematico.

È su questo sfondo che la logica assumerà un ruolo ancor più decisivo nell'indagine sui fondamenti, ponendosi problemi di tipo nuovo ed elaborando tecniche e metodi che non hanno parallelo in tutta la sua storia precedente.

L'occasione per questa rottura è costituita dalla crisi dei fondamenti, che almeno formalmente si aprì nel 1902 con la scoperta da parte di Russell della contraddittorietà della teoria cantoriana (e fregeana) degli insiemi.[1] Aveva così inizio una nuova fase dell'indagine sui fondamenti, caratterizzata dall'emergere di problemi e posizioni che solo in modo indiretto e mediatamente si ricollegano

[1] Il 1902 è l'anno in cui Russell, come sappiamo, comunicò a Frege l'antinomia scoperta nel sistema dei *Grundgesetze*; in questo senso esso segna la data «ufficiale» di apertura della crisi dei fondamenti, in quanto si è già detto che almeno dal 1895 (Cantor) e ancora nel 1897 (Burali-Forti) si era a conoscenza di antinomie nella teoria degli insiemi. Ovviamente, l'associare Frege (o anche Dedekind) a Cantor relativamente alla «teoria degli insiemi» ha solo un significato metaforico riferito alla presenza essenziale nelle rispettive teorie di tutti questi autori del concetto di insieme (estensione di un concetto, sistema).

alle ricerche del XIX secolo. Sono questi problemi e queste posizioni che costituiscono i punti di riferimento iniziali dell'indagine logica contemporanea.

Del paradosso di Russell e del suo significato per il lavoro di Frege abbiamo già avuto occasione di parlare; è giunto ora il momento di vedere in che senso esso portava (o riportava) alla luce interrogativi ineludibili per tutti. Si può dire che l'antinomia che Russell era riuscito a derivare dal V principio dei *Grundgesetze* mostrava inequivocabilmente che l'assunzione dell'esistenza, per ogni condizione linguistica, dell'insieme di tutti e soli gli oggetti che la soddisfano, portava a contraddizione ed era pertanto inaccettabile così com'era. In questo modo «vacillava sui suoi fondamenti» non solo la ricostruzione fregeana dei numeri naturali, ma, sembrava, la stessa teoria cantoriana degli insiemi, alla cui base si trova un principio di comprensione per la costruzione di insiemi analogo a quello fregeano. Con essa erano compromessi alcuni sviluppi matematici che, dalle prime applicazioni di Cantor allo studio degli insiemi di punti, erano stati il maggior titolo della teoria degli insiemi per un concreto riconoscimento da parte del mondo matematico. Se al I congresso internazionale dei matematici tenutosi a Zurigo nel 1897, Jacques Hadamard (1865-1963) e Adolf Hurwitz (1869-1919) avevano fornito prove concrete delle grandi possibilità che era lecito aspettarsi dalla teoria cantoriana per lo sviluppo dell'Analisi, al IV congresso, tenutosi a Roma nel 1908, Henri Poincaré (1845-1912) parlava ormai della *Mengenlehre* come di una malattia, «un meraviglioso caso patologico» che, ben lungi dal poter offrire aiuti al *corpus* della matematica, necessitava essa stessa di cure urgenti.

All'origine di questo mutamento di prospettiva non stava però la sola antinomia di Russell; altre se ne erano aggiunte, più gravi non tanto dal punto di vista formale, quanto per l'inconciliabilità delle soluzioni che sembravano concedere al problema della natura del procedere matematico. Il paradosso russelliano colpiva, diciamo così, il carattere troppo liberale ed illimitato della definizione cantoriana di insieme come estensione di proprietà arbitrarie. In questo senso esso era letale per il sistema fregeano in cui il concetto di proprietà non poteva essere limitato a meno di perdere decisamente il suo carattere logico generale, ma lo era molto meno per la pratica matematica allora corrente. Tutto sommato il mate-

matico non parlava mai, in concreto, di proprietà di insiemi arbitrari, ma sempre di insiemi di punti, di numeri ecc. di insiemi cioè appartenenti a un insieme già dato. Nella pratica, quindi, il paradosso di Russell aveva un riflesso limitato e poteva essere considerato come non pertinente. Lo stesso Cantor, che pure nei suoi progetti mirava a una teoria generale degli insiemi astratti e non solo degli *insiemi di* qualche cosa, degli insiemi concreti, non l'aveva considerato tutto sommato un ostacolo insormontabile, quando, prima ancora di Russell, ci si era imbattuto, ma l'aveva *positivamente* utilizzato.[2] L'insidia maggiore proveniva da altre ambiguità, più sottili, implicite nell'idea degli insiemi come estensioni di proprietà, così che il paradosso di Russell diveniva la spia di una fra le tante oscurità implicite nell'idea cantoriana e fregeana d'insieme, ma non era certo l'unica.

Questo non tardò a venire alla luce quando in una comunicazione del 1905 Julius König (1849-1913) presentò il suo famoso argomento sull'impossibilità di considerare la potenza del continuo come un alef. La dimostrazione come tale non è accettabile, ma una volta ridotta alle sue giuste proporzioni si trasformava in una nuova antinomia ben più resistente all'analisi di quella russelliana. Fatto oltremodo significativo, proprio nello stesso anno anche il francese Jules Richard (1862-1956) presentava l'antinomia che porta il suo nome. Entrambi i paradossi, quello di König e di quello Richard, si incentravano sullo stesso problema e sostanzialmente si possono considerare uno la variante dell'altro. Anche se, nella letteratura, fu l'antinomia di Richard che più spesso venne citata, è in quella di König che l'*impasse* risultava in forma più drastica e, almeno apparentemente, più disastrosa.

Il ragionamento di König è estremamente diretto. Si supponga vera l'ipotesi del continuo di Cantor; allora il numero cardinale di $\mathscr{P}(N)$,[3] l'insieme dei numeri reali, è un alef, quindi è bene ordinabile. Consideriamo ora la famiglia di tutti i sottoinsiemi di N com-

[2] Propriamente Cantor non si era imbattuto nel paradosso di tutte le classi, ma come sappiamo in quello del massimo numero cardinale (o del massimo numero ordinale) che come vedremo sono per così dire di una stessa categoria logica.

[3] Ricordiamo che l'ipotesi ristretta del continuo di Cantor consisteva nell'affermare l'uguaglianza del numero cardinale dell'insieme potenza di un insieme numerabile con l'alef successivo ad $\aleph_0$, ossia con $\aleph_1$; questo comportava il buon or-

presi in $\mathscr{P}$(N) *e definibili utilizzando un linguaggio basato su un alfabeto finito* (un alfabeto in altri termini analogo a quello di qualsiasi lingua naturale) che per ipotesi deve poter esprimere tutti i nostri concetti logici e matematici. Chiaramente, questa famiglia essendo per necessità numerabile, non potrà coincidere con tutto $\mathscr{P}$(N), che è più che numerabile per il teorema di Cantor. Prendiamo ora il primo elemento di $\mathscr{P}$(N) non incluso nella famiglia dei definibili; questo è possibile in quanto per ipotesi $\mathscr{P}$(N) è bene ordinato e quindi il complemento della famiglia dei definibili rispetto a $\mathscr{P}$(N) non è vuoto e ha un minimo. Sia *a* l'elemento in questione: in quanto primo degli elementi non definibili, *a* sarà non definibile; d'altra parte la locuzione «primo dei non definibili» è un'espressione del nostro linguaggio che definisce *a*: *a* è quindi definibile. Ci si trova così di fronte ad una contraddizione che pone in antagonismo i due pilastri della costruzione cantoriana: la concezione degli insiemi come estensioni di proprietà specificate da condizioni linguistiche da una parte, e dall'altra l'ipotesi del continuo come affermazione della possibilità di fondere in un'unica nozione (quella di alef) la considerazione cardinale e ordinale degli insiemi. Fatto ben più grave, l'argomento, diversamente da quello di Russell, non si riferisce a insiemi astratti individuati entro l'universo di tutti gli insiemi, ma ad insiemi concreti, parti di una collezione ben delimitata come quella dei numeri naturali, che costituisce addirittura il primo gradino della scala cantoriana del transfinito.

A questo punto, le alternative che sembravano aprirsi erano egualmente scoraggianti. Se si lasciava cadere la possibilità di ben ordinare gli insiemi ci si trovava di fronte a difficoltà insormontabili nello sviluppo delle parti più elementari della teoria del transfinito; in particolare, cadendo di conseguenza anche la confrontabilità dei cardinali, risultava praticamente impossibile risolvere anche i più semplici problemi di aritmetica cardinale. Questa via, in altri termini, se poteva salvare la teoria cantoriana dei cardinali e degli ordinali numerabili, vanificava la parte più ambiziosa della teoria del transfinito togliendole il suo principale strumento di-

dinamento, che Zermelo proprio nel 1904 avrebbe dimostrato introducendo proprio a questo scopo l'assioma di scelta. Qui N è al solito l'insieme dei numeri naturali; $\mathscr{P}$(N) è allora l'insieme di tutti i sottoinsiemi di N ossia, appunto, l'insieme di tutti i numeri reali.

mostrativo. D'altra parte, se si attribuiva al paradosso il significato di dimostrazione della sostanziale indeterminatezza del concetto generale di «insieme definito» ci si trovava in una situazione ancora più imbarazzante dal punto di vista logico. Negare infatti un significato determinato e univoco al concetto di «insieme definito» significava togliere determinatezza anche al principio cantoriano di considerare come dato ogni insieme individuato da una proprietà formulata linguisticamente. Come nel caso del paradosso di Russell, ma in una forma più sfuggente, ci si trovava di fronte alla necessità di negare un significato preciso al concetto di individuazione mediante una condizione linguisticamente definita, di negare significato, quindi, alla stessa nozione di estensione di proprietà.

Entrambi questi paradossi – ma del resto anche quelli più o meno contemporanei di Burali-Forti e di Richard – sembravano colpire le nozioni centrali, dal punto di vista matematico e logico, della teoria cantoriana ma quel che più conta, proprio quelle nozioni che erano direttamente coinvolte nella ricostruzione fregeana e dedekindiana delle strutture classiche dei numeri naturali e reali. Se ora teniamo conto del valore paradigmatico che queste ricostruzioni avevano avuto come esempi della possibilità di determinare (ricostruire) i concetti matematici fondamentali su basi puramente logiche o comunque libere da presupposti empirici o psicologici, ci si può rendere conto dell'effetto disastroso che i paradossi ebbero non sulla pratica matematica, che come abbiamo visto ne fu colpita solo minimamente, ma sulla *concezione* della matematica che i lavori di Frege, Dedekind e Cantor avevano portato avanti. La possibilità di creare concetti non era poi così «libera» come Cantor aveva supposto: dovevano esistere condizioni limitanti di cui sembrava difficile affermare la natura puramente logica.

Come ebbe ad osservare più tardi Gödel, i paradossi non costituivano infatti un problema insormontabile per la teoria degli insiemi, ma per la teoria delle proprietà e per ogni concezione della matematica (come quella logicista) che ad essa facesse riferimento. Le antinomie a questo riguardo sembravano non lasciare scampo: o si rinunciava a concepire come determinato in modo assoluto, puramente logico, il concetto di insieme come estensione di una proprietà, o si doveva ammettere che c'era qualcosa di pro-

fondamente sbagliato nella analisi logica che si era condotta. In entrambi i casi la questione riguardava ormai non più la sola teoria degli insiemi, ma la natura stessa della costituzione dei concetti logici. Di fronte a questa nuova situazione gli atteggiamenti furono diversi a seconda del ruolo che si assegnava alla teoria degli insiemi e del significato logico generale che si dava ai suoi concetti base.

Per i creatori dell'approccio logicista, come è naturale, l'effetto fu drammatico: Frege riconobbe il sostanziale fallimento del suo programma e cessò di fatto ogni attività in questo senso specifico; Dedekind, non meno compromesso di Frege, per lungo tempo si oppose alla ripubblicazione di *Essenza e significato dei numeri*, e abbandonò definitivamente l'impresa. All'estremo opposto la posizione di alcuni matematici, come ad esempio A. Schönflies (1853-1928) che ricorrendo a una netta distinzione fra matematica e logica riteneva di poter uscire dalle difficoltà relegando i paradossi nell'ambito della problematica puramente filosofica.

Accanto a queste posizioni, estreme in senso opposto, si situano i veri e propri tentativi di analisi delle antinomie. La storia del loro articolarsi e della problematica cui diedero origine coincide ormai con la storia della logica almeno per i primi trent'anni del nostro secolo, ed è compito dei prossimi paragrafi darne una presentazione non del tutto approssimativa. Qui vogliamo tuttavia tracciare un panorama generale di tale sviluppo, che ne indichi per sommi capi le grandi linee direttive e i loro collegamenti. Se questo ci porterà di necessità a istituire distinzioni e antitesi troppo drastiche potrà permettere però di avere una visione d'insieme del mutuo intrecciarsi di problematiche e posizioni che ad un'analisi più dettagliata potrebbero sembrare del tutto irrelate.

Come si è visto sopra, l'aspetto più rilevante dei paradossi era il fatto che essi colpivano la connessione stabilita da Frege, Cantor e Dedekind fra logica e matematica; in altri termini, veniva messa in crisi la possibilità di definire con un vocabolario e principi puramente logici, privi di riferimento a dati esterni, e quindi assoluti, i concetti matematici fondamentali. Chiarito questo, è anche ovvio che le posizioni che emergeranno in seguito alla scoperta delle antinomie vorranno tutte rispondere a una domanda di fondo: che rapporti esistono fra logica e matematica, quali fatti giustificano la significatività e la verità della matematica?

Cronologicamente, la prima risposta sufficientemente articolata, almeno a livello programmatico, è quella di un netto rifiuto di ogni possibile fondazione della matematica su basi logiche con la conseguente affermazione dell'autonomia dei concetti matematici fondamentali dalla logica. È questo, sostanzialmente, il punto di partenza di Poincaré, punto di partenza comune tanto della posizione «empirista» della scuola degli analisti di Parigi[4] quanto più tardi del ben più radicale intuizionismo brouweriano. Corollario immediato di questa posizione doveva poi essere l'individuazione di una «fonte» dei concetti matematici (almeno di quelli alla base della pratica matematica classica, teoria degli insiemi esclusa) che ne garantisse l'oggettività, o in altri termini la possibilità di studio rigoroso.

La risposta a questo problema, seppure con enfasi diverse, è comune tanto a Poincaré quanto a Luitzen Egbertus Jan Brouwer (1881-1966): la matematica ha un contenuto suo proprio che le proviene direttamente e senza mediazione dall'intuizione ed è come tale indipendente tanto dall'esperienza sensoriale quanto dalla strutturazione logica. In questo senso, la logica non è che una veste che per scopi di comunicazione viene *imposta* a contenuti che ne sono del tutto indipendenti. La logica rimane quindi una pura forma, irrimediabilmente vincolata a strutture linguistiche e come tale niente affatto normativa nei riguardi di contenuti che provengono da ben altra fonte, l'intuizione appunto. Scompare così il ruolo creativo della logica come capace di ricostruire e addirittura creare *ex novo* i concetti matematici di base. In questa prospettiva le antinomie sono le eloquenti illustrazioni di come, una volta affidatisi al puro gioco formale della logica e dell'apparato linguistico, ci si possa imbrogliare come in uno scioglilingua. Per Poincaré, i paradossi non colpiscono minimamente la matematica «vera» quella fondata su precisi dati intuitivi, proprio perché la nozione d'insieme cantoriana è una creazione puramente logico-linguistica sprovvista di contenuti intuitivi.

Il problema diviene allora quello di giustificare la pratica mate-

[4] I cui aderenti vengono anche detti «semiintuizionisti» riservando la qualifica di «intuizionisti» o «neointuzionisti» ai seguaci della dottrina di Brouwer. Fra i semiintuzionisti francesi della scuola di Parigi vanno ricordati Émile Borel (1871-1956), René Baire (1874-1932) e Henri Léon Lebesgue (1875-1941).

matica isolandone i contenuti intuitivi; solo a questo patto infatti, questa posizione «intuizionistica» (in senso lato) non si risolve in un puro e semplice nominalismo. Affermare che la matematica ha contenuti intuitivi è una cosa; individuare tali contenuti e giustificarne lo *status* è un'altra cosa. È proprio nella risposta a questa domanda fondamentale che le posizioni di Poincaré, della scuola di Parigi e di Brouwer si distanziano tra loro e dalle tesi di Kronecker, storicamente contemporanee alla creazione di Cantor. Alla base di tutte queste concezioni sta un unico concetto, pur diversamente caratterizzato: quello di *costruzione* (matematica); ed è sulla sua base che si articola l'intera analisi del contenuto intuitivo della matematica.

Come infatti per Frege e per Cantor esisteva il problema di *definire* i vari enti matematici partendo da semplici dati della logica, così per queste forme di costruttivismo e intuizionismo il problema centrale è quello di come *costruire* a partire da dati intuitivi i concetti base della matematica classica. È a questo livello che le posizioni divergono; pur accettando tutti come dato l'esistenza della successione dei naturali concepita come possibilità in potenza di iterare la stessa operazione di passaggio al successore, diverse sono le posizioni sul come, da questa totalità infinita potenziale, si possano ottenere altri enti, altre totalità. In questo caso è dall'analisi diretta dell'attività intuitiva che scaturiscono le risposte al problema.

Come osserva Poincaré, se si concepisce la creazione degli enti matematici come correlato dell'attività intuitiva risulta evidente che l'unico modo per ottenere nuovi enti da enti dati è quello di definire questi ultimi in modo *contenutisticamente* sensato a partire dai primi. In altri termini, è possibile ottenere oggetti matematici solo ricorrendo ad enti già dati senza far riferimento a totalità non ancora costruite, quindi in particolare alla totalità cui gli enti definitivi devono appartenere. È questo quel *principio del circolo vizioso* che costituisce il punto di partenza della posizione cosiddetta *predicativista*. Su questa base, Poincaré, come gli analisti della scuola di Parigi e lo stesso Brouwer basano la loro critica alla teoria cantoriana e ai tentativi (di Russell come di Zermelo) di porre rimedio alle antinomie. Una volta accertato, proprio mediante l'insorgere dei paradossi, che l'unico modo per fondare al di fuori di ogni dubbio la matematica è quello di riportarsi all'analisi dell'attività

matematica che è alla sua base, e una volta riconosciuto che l'intuizione pone come dati dei contenuti solo se è in grado di afferrarne le modalità di costruzione, risulterà chiaro che ogni riferimento alla totalità degli enti matematici, o anche dei soli insiemi, era destituito di ogni senso intuitivo; si riduceva ad un semplice gioco linguistico il quale anche se, opportunamente regolato, poteva non portare ad antinomie risultava comunque privo di contenuto.

Alla operazione di potenza cantoriana e al suo principio di comprensione, Poincaré opponeva così il principio del circolo vizioso come criterio di significatività, di specificazione d'insiemi. Se l'analisi di Poincaré, pur nella sua fecondità come strumento di individuazione dell'origine delle antinomie, sembrò ai non costruttivisti come Bertrand Russell e Ernst Zermelo (1871-1953) (e allo stesso König) una limitazione troppo pesante per la definizione degli insiemi, per i costruttivisti in senso proprio come Brouwer non poteva, viceversa, essere ancora sufficiente.

La critica di Poincaré, infatti, in ultima istanza, aveva colpito solo un aspetto del realismo e «logicismo» cantoriano, e cioè la sua pretesa di affermare l'esistenza di enti (o insiemi) sulla semplice base della loro definibilità (e non costruibilità) a partire da enti già dati. La sostituzione del concetto di definibilità, puramente formale, con quello basato sull'analisi intuitiva di costruzioni partendo dal principio del circolo vizioso, aveva eliminato proprio questa assunzione, mostrandone il significato puramente linguistico. Nell'analisi di come si costruiscono e si realizzano le classificazioni su cui si fonda il ragionamento logico – che per Poincaré si riduce essenzialmente alla sillogistica o, nei termini che userà poi Brouwer, alla matematica del tutto e della parte – Poincaré, secondo Brouwer, si era pur sempre fondato su considerazioni linguistiche e logico-formali; egli aveva infatti ammesso che un insieme (o un ente qualsiasi) definibile «in modo finito» esclusivamente in termini di enti già dati poteva essere considerato come dato, come intuitivamente comprensibile. Era proprio questo, a parere di Brouwer, il limite del predicativismo di Poincaré, questa sostanziale accettazione di una nozione ancora «linguistica» e formale, non sorretta, o meglio, sostituita da un'analisi *diretta* dell'effettiva attività intuitiva di costruzione a partire da dati intuitivi sicuri. In questo modo Poincaré, come anche Borel e Baire, aveva-

no potuto accettare come fondata, anche se con esitazioni e tentennamenti, larga parte della teoria classica del continuo.

Pur ricollegandosi esplicitamente al loro lavoro, Brouwer, ad iniziare dagli anni attorno al 1910, tentò appunto l'impresa più audace di ricostruire la matematica *ex novo*, basandosi su un'analisi diretta del concetto di costruzione e rigettando così ogni forma di compromissione con la matematica classica o con l'atteggiamento logicista. Nella sua lunga vita e con l'aiuto di numerosi collaboratori, Brouwer riuscì così a creare una vera e propria matematica il cui significato va cercato proprio nella fondazione intuitiva e non puramente linguistico-formale che si trova alla sua base. Il salto dalla posizione di Poincaré è così molto netto: Poincaré e gli stessi semiintuizionisti francesi si erano posti il problema di ricostruire la maggior parte possibile della matematica classica basandosi sul criterio della sensatezza contenutistica (ma abbiamo visto con che limitazioni) dei suoi concetti base; per Brouwer il problema è quello di creare una nuova matematica, abbattendo i «rami secchi» nati da puri giochi linguistici privi di contenuto intuitivo genuinamente costruttivo. Cade così la teoria cantoriana degli insiemi, cade la teoria stessa del continuo classico. Al loro posto, in un lungo giro d'anni, Brouwer riuscì a sostituire altri concetti base, quali quelli di spiegamento, di specie, di successione di libera scelta, che permettono di costruire una vera e propria teoria intuizionista del continuo.

Alla base di questi concetti, come abbiamo detto, sta la posizione fondamentale dell'intuizionismo di netto rifiuto di ogni contaminazione e imposizione da parte dell'apparato logico e linguistico. Questo non significa ovviamente che Brouwer non usi strumenti dimostrativi, ché ciò sarebbe negare la natura esatta del pensiero matematico; ma afferma recisamente che la logica non ha una legittimità sua propria che sia superiore a quella data dai contenuti matematici specifici. La logica classica di cui ancora si era fidato Poincaré è un'estrapolazione puramente linguistica da situazioni specifiche, da metodi dimostrativi validi concretamente in ambiti più ristretti. Brouwer così rigetta il principio distintivo della logica classica, il terzo escluso, appunto come un'illecita estrapolazione non valida in generale. Se infatti l'uso del terzo escluso è lecito su domini finiti in quanto in questo caso non fa che riflettere un atto intuitivo ben preciso (la possibilità cioè di co-

struire un controesempio di un'affermazione non vera sul dominio in questione) il suo impiego non è invece ammissibile in un dominio infinito o in generale in presenza di proprietà non dominabili intuitivamente e non effettivamente decidibili. Con questo, la critica brouweriana raggiunge il suo punto più decisivo: il rifiuto della logica classica stessa come contenutisticamente non giustificata. Non si può immaginare contrapposizione più netta con l'atteggiamento logicista. Se infatti per Frege la logica classica era al di fuori di ogni dubbio, forma in grado di costituire gli stessi contenuti matematici, per Brouwer essa non solo risulta subordinata e secondaria rispetto all'operare matematico concreto, ma addirittura non accettabile.

Dalla crisi dei fondamenti concepita come crisi del rapporto tra legittimità logica e matematica, emerge così una posizione che, con una chiarezza e un'articolazione sino ad allora mai viste, poneva per la prima volta il problema della validità stessa delle tradizionali leggi logiche. Soprattutto dopo la codificazione dei principi della logica e dell'aritmetica intuizioniste ad opera di Arend Heyting (1898-1980) verso il 1930, la posizione di Brouwer diventerà una delle correnti centrali della ricerca sui fondamenti della matematica di questo secolo. In senso lato si può affermare che fu sotto lo stimolo della sua critica che si giunse alla chiarificazione di concetti quali quello di computabilità, costruzione ecc., allo studio di logiche più deboli di quella classica, ed in generale all'interesse sempre più articolato per il contenuto costruttivo delle singole teorie matematiche. Non è esagerato dire, in altri termini, che proprio alla critica intuizionistica si deve in gran parte l'estendersi e l'approfondirsi dell'analisi iniziata con i paradossi. Lo stesso formalismo, come pure il logicismo, come vedremo, pur partendo da posizioni nettamente diverse, ricevettero potenti stimoli alla propria autochiarificazione dal confronto diretto col predicativismo prima e con l'intuizionismo poi. Ed è proprio da un tentativo di riprendere in chiave logicista il principio del circolo vizioso di Poincaré che prese appunto le mosse Bertrand Russell in collaborazione con Alfred North Whitehead[5] nel suo tentativo di costrui-

[5] Pur essendo opera di collaborazione, i *Principia mathematica* devono in massima parte a Russell il loro impianto «filosofico» logicista nell'accezione qui intesa. Questa conclusione può ricavarsi e dalle dichiarazioni dei due autori e dall'impostazione e dal contenuto dei lavori di Russell anteriori ai *Principia* che verranno

re su basi logiche l'intero edificio della matematica classica, teoria degli insiemi inclusa.

I *Principia mathematica*, la cui prima edizione risale al 1910-13, nascono infatti all'insegna di questa doppia influenza: da una parte l'analisi fregeana della logica e dei fondamenti dell'aritmetica, dall'altra i dibattiti suscitati dal predicativismo di Poincaré e più in generale dai paradossi. Il logicismo di Russell si pone come punto di confluenza di questi due filoni di ricerca; più radicale per alcuni versi di quello fregeano, esso presenta nel contempo problemi e aspetti tecnici che ne fecero ben presto un punto di riferimento obbligato nell'indagine dei fondamenti non solo per i logicisti, ma per tutti coloro per i quali la questione dei rapporti fra matematica e logica non poteva risolversi così drasticamente come era stato fatto da Poincaré e dai semiintuizionisti.

L'obiettivo centrale dei *Principia* è infatti quello di mostrare come, una volta analizzato opportunamente il meccanismo delle antinomie, fosse possibile ricostruire l'edificio della matematica classica a partire da nozioni e principi logici. È questo il logicismo russelliano. Se da una parte esso può considerarsi la ripresa, poste le dovute precauzioni contro le antinomie, del programma fregeano, dall'altra costituisce un fatto decisamente nuovo, alla cui origine stanno elementi e stimoli eterogenei. Per Russell, diversamente che per Frege, non esistono eccezioni alla riduzione della matematica alla logica: teoria dei numeri, Analisi, teoria del transfinito, geometria, sono tutte in egual grado diverse articolazioni di un contenuto logico unico, diverse organizzazioni (o esplicitazioni) di quelle «forme» che costituiscono l'oggetto ed il contenuto stesso della logica. In questo senso il logicismo russelliano na-

citati nel paragrafo 2.1; Whitehead concorse attivamente e pesantemente al lavoro comune soprattutto per quanto riguarda la parte più propriamente matematica. Va ricordato tuttavia che già nel 1898 Whitehead aveva pubblicato quel *Trattato di algebra universale con applicazioni* che, secondo Couturat che lo recensì nel 1900, portava alla «conclusione filosofica» che tutta la matematica è fondata sull'algebra della logica. Come già detto tuttavia (cap. II, par. 7) questo trattato di Whitehead va riguardato piuttosto come una prima sistematica presentazione della matematica intesa come «sviluppo di tutti i tipi di ragionamento formale, necessario e deduttivo», nella quale Whitehead sembra esplicitamente trascurare la «dimensione» filosofica, salvo eventualmente che nella parte introduttiva. Ecco perché noi, parlando in generale di logicismo, ci riferiremo essenzialmente a Russell. Naturalmente ciò non significa che Whitehead non abbia sviluppato una sua propria «filosofia».

sce come esplicito bilancio ed interpretazione di *tutto* il lavoro sui fondamenti fatto nel secolo precedente – indagini sui fondamenti della geometria incluse – e si pone quindi in modo naturale come antagonista delle nuove concezioni predicativiste e semiintuizioniste che partivano da ben altra analisi di quello stesso lavoro.

L'assunto di fondo, per Russell, è infatti il carattere di organizzazione formale e di schematizzazione di contenuti extralogici di ogni concetto e teoria matematica. Ben lungi quindi dal trovare il proprio significato nel collegamento con specifici contenuti matematici (intuizione della successione dei naturali ecc.) le singole teorie matematiche trovano la loro ragion d'essere e la loro determinatezza nel tipo di schemi generali astratti che individuano. Il loro ruolo nel sistema della conoscenza umana è appunto quello di fornire e rendere esplicite queste forme. Che poi queste, per ragioni legate alla natura particolare del mondo della nostra esperienza o intuizione, risultino applicabili oppure no, di volta in volta, ad ambiti specifici e limitati, può spiegare il nostro interesse o meno per esse, ma non toglie nulla al loro carattere puramente logico e astratto: la natura di scienza esatta della matematica discende da questo loro carattere, e non dalla connessione con particolari ambiti intuitivi. Dire però che le singole teorie matematiche non hanno un contenuto specifico offerto dall'intuizione o dall'esperienza non significa per Russell che l'intero edificio della matematica non abbia contenuto e sia un puro sistema formale. Sono le stesse forme, isolate e definite in termini logici generali, che costituiscono il contenuto della matematica nel suo complesso, ed è in questo senso che la matematica non è che un aspetto della logica e su di essa si fonda. L'obiettivo, il contenuto della logica, è così costituito dalla totalità delle forme e degli schemi considerati nella loro universalità, di cui il matematico fa uso in ambiti determinati: i teoremi della logica (e della matematica) sono tutti gli enunciati universali, veri, riguardanti le sole costanti logiche fondamentali; sono forme pure, svincolate da ogni riferimento particolare.

In questo senso, il ribaltamento rispetto alle concezioni semiintuizioniste è completo: la logica ritorna al suo posto centrale, riconfigurandosi come in Frege, ma in maniera ben più determinata, come costitutiva degli oggetti stessi della matematica.

Se i numeri naturali vengono definiti come in Frege e costituiscono il materiale di base da cui si ottengono successivamente gli

altri oggetti fondamentali dell'Analisi, la geometria per Russell si può ricondurre alla logica una volta che si vedano i suoi enunciati come enunciati condizionali: supposto che $\mathcal{A}$ sia la congiunzione, poniamo, degli assiomi della geometria euclidea del piano, il problema di dimostrare l'enunciato $\mathcal{B}$ diviene quello di provare, all'interno della Grande Logica, l'implicazione $\mathcal{A} \rightarrow \mathcal{B}$. In questo senso gli assiomi della geometria non fanno che fissare forme possibili, all'interno dell'universo logico, relativamente alle quali gli ordinari enunciati geometrici vanno interpretati.

Ma che cosa è la logica? È questa la domanda fondamentale a cui Russell, e dopo di lui tutto il logicismo post-fregeano, era costretto a dare una risposta per salvare il senso del programma generale. Per Frege la logica si poteva configurare come un edificio compatto e sostanzialmente unitario fondato da una parte sulla esplicitazione dei canoni di inferenza, dall'altra su principi (come il principio V dei *Grundgesetze*) che stabilivano le condizioni necessarie e sufficienti affinché un concetto, specificato facendo ricorso al solo significato delle costanti logiche, potesse considerarsi determinato e quindi possibile oggetto di predicazione. In questo modo tutti i possibili concetti astratti, purché determinati, e quindi tutte le possibili forme, venivano ad essere costrutti pienamente dominabili una volta che si fossero stabiliti i significati delle costanti logiche fondamentali: era questa la Grande Logica. Con la scoperta delle antinomie però risultava chiaro che ciò non era possibile e che solo apparentemente questo edificio logico era così netto nelle sue articolazioni fondamentali.

Le alternative che si aprivano erano due, come avevano chiaramente sottolineato Poincaré e i semiintuizionisti: o si rinunciava all'idea della matematica come studio delle forme e quindi al programma logicista, oppure si dichiarava inaccettabile la logica usata da Frege e ci si sobbarcava l'onere di costruirne una plausibile e adeguata al programma. Il grande contributo di Russell sta nell'aver mostrato come la seconda alternativa fosse praticamente realizzabile, ridando così significato al programma logicista.

Non deve stupire allora che sia proprio dal principio del circolo vizioso di Poincaré che prende le mosse la nuova analisi russelliana e la sua teoria dei tipi ramificata. Questa teoria, che come vedremo Russell adottò nei *Principia* dopo tante esitazioni, non significa affatto una rinuncia al programma ed una conversione a una

visione basata sul sintetico a priori o sull'intuizione, quanto la presa di coscienza della necessità di analizzare *all'interno* gli stessi concetti di proposizione, di universalità ecc. che Frege aveva assunto come inanalizzati. È il primo passo di quella ricostruzione globale dei concetti logici fondamentali che il logicismo avrebbe posto come suo obiettivo.

L'atteggiamento con cui Russell si pone di fronte al problema è infatti decisamente «sperimentale»; in conformità alla sua visione generale della logica (e della matematica) come scienza delle forme astratte definite ricorrendo a un dato vocabolario di idee logiche primitive e sulla base di pochi principi logici fondamentali, egli si impone come metodo esclusivo quello di passare in rassegna il dominio sterminato di concetti, schemi definitori e dimostrativi in pratica usati nell'intera matematica, di analizzarne l'articolazione interna, per giungere infine all'isolamento delle idee e dei principi fondamentali. Questo gli permette di raggiungere due obiettivi: da una parte l'individuazione dei concetti base in grado di evitare le antinomie, dall'altra la prima vera unificazione e ricostruzione della matematica sulla base di un esiguo numero di concetti e principi logici fondamentali.

Sfondo di tutto questo enorme lavoro sono i contributi dell'analisi logica del secolo XIX, non solo le ricerche di Frege, Dedekind e Cantor, ma tutta la tradizione legata alle indagini sui fondamenti della geometria: von Staudt, Pasch e la scuola italiana, di cui abbiamo parlato nel capitolo precedente. Accanto a queste, il complesso delle ricerche di tipo algebrico (algebra della logica, quaternioni, vettori, ecc.) una cui sistemazione Whitehead aveva già offerto nel suo *Treatise*. Tra le parti più vitali – anche se meno lette oggi – dei *Principia* sono le sezioni che si ispirano a questa tradizione e che più tardi avrebbero influenzato – soprattutto in Polonia come abbiamo visto – lo sviluppo astratto di algebra, topologia, Analisi funzionale, ecc. Detto questo, ciò che va sottolineato con enfasi è che mai prima dei *Principia* si era raggiunta la *concreta* dimostrazione ostensiva di come di fatto poche idee primitive e pochi assiomi potessero permettere la ricostruzione di tutta quanta la matematica classica, teoria del transfinito e geometria incluse. In questo senso, meriti e demeriti intrinseci della teoria dei tipi a parte, i *Principia mathematica* costituiscono una pietra miliare nella storia della logica di questo secolo. Nei tre volumi di cui sono co-

stituiti, l'unificazione di tutta la matematica classica sulla base di alcuni assiomi e nozioni fondamentali (siano poi queste puramente logiche o no) non è affermata a parole, ma *realizzata in concreto.*

Questo aspetto non va dimenticato se si vuole afferrare il senso che il logicismo russelliano (e post-russelliano) ha avuto nella storia di questo secolo. L'obiettivo è quello di realizzare la ricostruzione di tutta la matematica, di unificarla sulla base di pochi concetti primitivi costituendo una Grande Logica come fondazione del pensiero astratto nella sua globalità. Di fronte a questa meta finale, le singole soluzioni adottate (teoria dei tipi ramificata, teoria dei tipi semplici ecc.) hanno il significato di tentativi di analisi, di esperimento, non di riduzione decisiva. Diversamente dai neointuizionisti e poi dai formalisti, il logicismo, quello di Russell soprattutto, non parte da assunzioni di fondo che poi cerca di rendere esplicite o di realizzare, ma si pone piuttosto come un lavoro di approssimazione, di adeguazione attraverso il raffinamento dell'analisi; e questo spiega le sue ambiguità di fondo e le sue incertezze.

La teoria dei tipi ramificata presentata da Russell e Whitehead nei *Principia* è l'esempio più significativo di questo atteggiamento. Non diversamente da Poincaré, Russell, come si è detto, identificava nelle definizioni impredicative l'origine delle antinomie e assumeva come criterio di significatività degli enunciati il principio del circolo vizioso. Solo che, diversamente da Poincaré, non concepiva questo principio come giustificato dalla necessità di sostituire, alle pure definizioni, costruzioni che per loro stessa essenza debbono essere predicativamente definite, ma lo vedeva come un *criterio generale che fissa i limiti di significatività degli enunciati*, o meglio *alcuni* di questi limiti: quelli sufficienti per evitare i paradossi. Alla necessità di concepire le funzioni enunciative come disposte in una doppia gerarchia di tipi e di ordini, Russell giungeva pertanto dall'analisi dei domini di significatività delle funzioni enunciative, anche quest'ultima condotta non sulla base di una nozione generale del concetto di campo di significatività, ma a partire dalla semplice constatazione «empirica» che, in qualunque modo si concepisca il significato di una forma enunciativa, per evitare i paradossi basta supporre che i domini di significatività siano disposti in una gerarchia, cosicché gli enunciati (o gli enti) paradossali vengono eliminati come non significanti.

Certo Russell è convinto che l'idea di fondo della teoria dei tipi,

di disporre tutti gli enti logici (funzioni enunciative, enunciati ecc.) in una scala secondo il grado di complessità, è un'idea naturale, intuitiva: l'universo risulta così concepito come disposto in una gerarchia il cui gradino più basso è costituito dagli enti logicamente semplici, inanalizzati, e ogni altro gradino è ottenuto dai precedenti collezionando e facendo riferimento alle totalità dei livelli più bassi. D'altra parte, però, ben lungi dal concepire questo quadro intuitivamente abbastanza plausibile come una descrizione adeguata della situazione, egli è perfettamente consapevole che l'ambiguità tipica delle nozioni logiche fondamentali (vero, falso, universalità, totalità ecc.) porta a non poter ricostruire concetti ed enunciati validi di uso essenziale. Come vedremo più in dettaglio nel par. 2.1, diversamente da Poincaré, Russell, pur essendo il primo a costruire una teoria di tipo predicativista (perché tale è la teoria dei tipi dei *Principia* una volta che non si postuli la riducibilità) non è affatto dell'idea che solo i ragionamenti e le definizioni predicativisti siano validi. Tutta la matematica classica, alla cui ricostruzione egli mira, fa uso essenziale di procedimenti non predicativi. Il contrasto tra queste due posizioni – predicativismo dell'apparato logico e non predicativismo della matematica che si vuole ricostruire – viene risolto da Russell postulando l'assioma di riducibilità che afferma la possibilità di trovare per ogni funzione enunciativa una funzione predicativa equivalente, in modo da poter dare un senso anche a quegli enunciati che predicativi non sono: è quindi possibile riprodurre entro la teoria dei tipi così allargata tutti i ragionamenti non predicativi validi nella matematica classica.

Per strano che possa apparire, è solo un appello alla sua necessità in questo recupero che Russell porta a difesa dell'assioma: è lo stesso tipo di giustificazione che usa pure per gli altri assiomi che è costretto a postulare, quello dell'infinito («esistono infiniti individui distinti») assunto per poter definire i numeri naturali, e quello della scelta («assioma moltiplicativo»). Nessuno di questi assiomi, per quanto intuitivamente plausibile, ha alcuna cogenza di logica necessità eppure Russell non si sente costretto a rinunciare alla sua tesi fondamentale per cui la matematica è logica. A tutte le obiezioni la sua risposta è sempre la stessa: questi assiomi sono presentati come risultato di un'analisi che potrà in futuro risultare troppo rozza, ma posseggono un grado induttivo di plausibilità

molto alto. Non c'è nulla di strano in questo per Russell: il dominio dei concetti logici non ci è immediatamente noto; la logica che li studia è quindi in buona parte una scienza che li costruisce mediante un lavoro di analisi della realtà, delle intuizioni concettuali in uso, del tutto simile a quelle del fisico che precede con approssimazioni induttivamente giustificate. L'essenziale è che l'analisi e i concetti «teorici» cui essa porta siano in grado di spiegare la realtà data, di ricostruirla. Il resto, come dirà il più grande logicista dopo Russell, Frank Plumpton Ramsey (1903-30), è questione di «dettaglio», di perfezionamenti successivi che non possono vanificare la tesi di fondo.

In questa posizione c'è una forte dose di ambiguità, inevitabile d'altra parte dopo il crollo del programma logicista nella forma fregeana ad opera dei paradossi. La necessità di ridefinire l'idea stessa di logica portava come conseguenza oscillazioni dovute alle esigenze contrastanti di non assumere una logica troppo ricca da ingenerare paradossi e di averla però sufficientemente potente da permettere una descrizione della matematica classica. La soluzione russelliana era un esempio di questa ambiguità, come fu ben presto realizzato da più parti. Nel suo tentativo di ricostruire la matematica su base logica Russell, come abbiamo visto, era stato costretto a postulare fra i suoi assiomi il principio di riducibilità, l'assioma di scelta e quello dell'infinito. Tutti questi assiomi in un modo o nell'altro facevano assunzioni il cui carattere esistenziale trascendeva decisamente ogni norma di verità logica plausibile e si ponevano così come innegabili assiomi di carattere extralogico o sintetico, come non tardò ad obiettare lo stesso Poincaré. La fondazione russelliana, in altri termini, poteva salvare l'idea logicista di una Grande Logica, come edificio unitario in cui ricostruire tutta la matematica, solo a patto di estendere la nozione di logica sino a comprendere principi di carattere sintetico. In che senso si poteva parlare ancora di logica con simili assunzioni? Lo stato logico di questi assiomi risultava nettamente problematico, e lo sforzo del logicismo post-russelliano fu proprio quello di mostrare come, mediante un'analisi ulteriore della base logica e del meccanismo dei paradossi, fosse possibile o chiarirne lo *status* logico o, più decisamente, eliminarli. Le tappe di questa ricerca di una precisazione del concetto di logica segnarono tutte, in un modo o nell'altro, lo sfaldamento dell'edificio unitario dei *Principia*.

Il punto centrale della controversia rimarrà pur sempre l'assioma di riducibilità, assioma che diversamente da quello dell'infinito o della scelta sembrava da una parte giustificato più che altro da considerazioni *ad hoc*, non motivate da un'analisi diretta dei concetti fondamentali, dall'altra introduceva nel sistema dei *Principia* un'assunzione esplicitamente esistenziale, un'ipotesi del tutto ingiustificata a priori, proprio sulla natura delle funzioni proposizionali. È significativo che su questo assioma, che costituisce il *trait d'union* tra una concezione realista ed una predicativista, si siano concentrati gli interessi dei logicisti dopo i *Principia*. Le alternative a proposito dell'assioma riaprivano così una dicotomia di fatto non risolta tra atteggiamento realistico e analisi costruttivistica.

Il primo passo verso questa separazione venne compiuto dal polacco Leon Chwistek (1884-1944) in alcuni lavori del 1923-24. In essi, rigettando decisamente l'assioma di riducibilità, Chwistek faceva emergere dai *Principia* una teoria dei tipi costruttiva che si poneva come precisazione delle idee predicativiste di Poincaré. In questo modo l'atteggiamento predicativista prendeva il sopravvento e il logico polacco era così costretto a rinunciare a parte della matematica classica la cui ricostruzione, senza riducibilità, risultava impossibile. Chwistek andava più in là, sostenendo la possibilità di derivare dal principio di riducibilità il paradosso di Richard. La dimostrazione, come fu sottolineato tra gli altri da Ramsey, era erronea, però metteva in luce un'ambiguità sottile presente nel sistema: in che senso l'assioma di riducibilità, che annullava «a meno di equivalenze» la distinzione tra funzioni predicative o meno, non vanificava proprio quell'applicazione del principio del circolo vizioso che permetteva di risolvere le antinomie?

Il fatto centrale, messo in luce da Ramsey nel 1926, stava in questo: le antinomie possono classificarsi in due categorie, quelle riguardanti nozioni matematiche (antinomia di Russell, di Burali-Forti ecc.) e quelle riguardanti nozioni logiche e linguistiche generali (antinomia di Richard, di König ecc.). Queste ultime hanno la proprietà di fare riferimento a contesti non estensionali come invece fanno le prime. In altri termini, non è possibile date due funzioni proposizionali estensionalmente equivalenti (tali cioè che per argomenti uguali abbiano valori uguali) e affermato che una delle due è definita in un dato modo, affermare che anche la se-

conda lo è. Nozioni come quella di definibilità (su cui si fonda il paradosso di Richard, come quello di König) non ammettono la sostituibilità di funzioni estensionalmente equivalenti: proprio perché esse fanno riferimento alle condizioni linguistiche con cui le funzioni sono specificate, la semplice equivalenza estensionale non è sufficiente per avere la sostituibilità in questi contesti. In questo senso il principio di riducibilità non implica paradossi come quello di Richard. Esso postula la possibilità di avere per ogni funzione non predicativa una predicativa estensionalmente equivalente; ma per derivarne un paradosso è necessario che l'equivalenza sia più stretta, e riguardi il modo di definizione, cosa questa che l'assioma di riducibilità non fa. Se questo mostra come il principio non sia contraddittorio, pone però in luce una limitazione di altro tipo. La Grande Logica di Russell, per poter ricostruire l'intero edificio della matematica classica, è costretta a espungere come irrilevanti distinzioni fondamentali, riguardanti nozioni come quella di definibilità, che fanno centro attorno alla intensione dei concetti. Come ammette lo stesso Ramsey, l'eliminazione della considerazione di queste distinzioni (con la conseguente possibilità di poter superare le antinomie del secondo tipo) trasforma la logica di Russell in un «sistema simbolico» che non può più aspirare a presentarsi, come avveniva per quello di Frege, quale «analisi del pensiero» (non psicologicamente inteso).

È accettando questa rinuncia che Ramsey nel 1925 giunge finalmente a presentare un sistema che, rigettando l'assioma di riducibilità, riesce d'altra parte a fondare l'intera matematica classica. Il sistema di Ramsey è in certo senso l'opposto di quello di Chwistek. Mentre quest'ultimo considerava come rilevanti le distinzioni intensionali e per questo rifiutava la riducibilità mantenendo però le distinzioni tra ordini delle funzioni proposizionali, Ramsey elimina anch'egli la riducibilità, rifiutando però d'altra parte anche la distinzione in ordini. La grande scoperta di Ramsey è infatti questa. Una volta considerate come non pertinenti a un sistema logico, e quindi non da ricostruirsi in esso, nozioni intensionali quali quelle coinvolte nei paradossi di secondo tipo, rimangono da superare solo i paradossi del primo tipo. Ma per fare questo non occorre gerarchizzare le funzioni proposizionali in ordini e tipi, basta limitarsi a questi ultimi. L'universo viene quindi stratificato solo in base alla complessità ontologica, «estensionale» delle fun-

zioni, e non secondo la complessità intensionale delle loro definizioni. La teoria dei tipi semplici che così emerge riesce a raggiungere i due obiettivi che Russell si era posto: l'eliminazione dei paradossi e la fondazione della matematica classica. A prezzo però di una rinuncia: quella di considerare come non degne di studio da parte della logica le distinzioni intensionali, rinunciando a porsi come plausibile teoria dell'universo delle proposizioni. In Russell considerazioni di questo tipo si presentavano attraverso la gerarchia degli ordini, ma venivano sostanzialmente lasciate da parte (via riducibilità) una volta che si trattava di fondare la matematica. Ramsey accentua così questo aspetto della «pratica» di Russell, eliminando *tout court* la distinzione. Ma è lecito espungere dalla logica simili considerazioni? L'altra via, quella di Chwistek, che questa rinuncia non richiede, sembrava d'altra parte disperata in quanto imponeva la rinuncia alla matematica classica e Chwistek formula anch'egli un sistema di tipi semplici nel 1926.

Se il sistema di Ramsey non presentava questa difficoltà ne presentava quindi un'altra, che riguardava la nozione stessa di logica, che sempre più si andava discostando dal progetto fregeano di Grande Logica, come sistema generatore di tutti i concetti astratti. Con Ramsey il logicismo giungeva così a una *impasse* che ne segna in certo senso la fine come forza propulsiva nell'indagine sui fondamenti. Con eccezionale lucidità Ramsey esamina una per una le possibilità che rimangono aperte. Abbiamo già parlato della riduzione estensionale e quindi realistica che la sua eliminazione degli ordini aveva portato. Ma un altro problema generale si prospetta. Per Russell, abbiamo visto, la logica era l'insieme di tutti gli enunciati universali veri riguardanti solo costanti logiche; ma in che senso formulazioni del genere erano legittimamente concepibili come leggi logiche? Questa soluzione rientrava nell'idea russelliana di logica come studio di tutti i possibili schemi astratti, ma un'idea del genere non risultava più accettabile una volta che si concepiva la logica non più come Grande Logica ma come sistema simbolico la cui natura logica doveva risultare da altre considerazioni. Sulla scia di Ludwig Wittgenstein (1889-1951) Ramsey è quindi portato alla identificazione della logica con l'insieme delle leggi riguardanti le costanti logiche non semplicemente vere, ma tautologiche. Questo era l'esito naturale di una ricerca della natura della logica che gradatamente, a partire dallo stesso

Russell, era stata costretta a restringere sempre più il suo orizzonte. A mano a mano, i principi esistenziali erano stati indeboliti e la logica, la Grande Logica, tendeva sempre più a restringersi a pura teoria inferenziale complementata da uno schema di comprensione vincolato dalle restrizioni sui tipi. Entro questa prospettiva gli altri assiomi critici di carattere esistenziale vennero a perdere sempre più plausibilità. Così, era costretto ad ammettere Ramsey, accadeva per l'assioma dell'infinito, la cui natura era decisamente non tautologica, così per l'assioma di scelta, che solo attraverso una drastica reinterpretazione poteva risultare accettabile.

Nel suo saggio del 1926 Ramsey concludeva con una dichiarazione di sostanziale scacco e se anche dopo di lui altre voci si alzarono per richiamare il programma fu solo, per così dire, in forma debole, come appello generale (e talora generico) alla eventuale possibilità di considerare la matematica come articolazione di un linguaggio logico generale, come un sistema organizzato sulla base di metodi dimostrativi puramente logici. Il tentativo di giustificare gli stessi assiomi come leggi logiche però scompare. Se anche più tardi si intrapresero, prima con Rudolf Carnap (1891-1970) poi con Willard Van Orman Quine ad esempio, tentativi di costruire sistemi logici generali, la tesi logicista della natura analitica o più in generale logica degli enti matematici sempre più andò perdendo sostenitori. Altre posizioni emersero, da una parte il costruttivismo e in particolare l'intuizionismo, ma soprattutto il formalismo che proprio negli anni in cui Ramsey scriveva andava prendendo forma definitiva.

Da più punti di vista il formalismo, nella formulazione che Hilbert ne diede negli anni venti, si può considerare l'erede, in senso del tutto particolare, del progetto logicista di una giustificazione della matematica classica, teoria degli insiemi inclusa. Comune è la convinzione che alla base dell'articolazione teorica di ogni disciplina matematica debba essere la logica classica nella sua forma piena, senza restrizioni imposte da considerazioni di carattere costruttivo; comune è la rivendicazione della possibilità di fondare l'intera matematica storicamente data su basi indubitabili e certe; comune infine è la valutazione della teoria cantoriana come culmine della matematica classica, al contempo ramo specifico e sfondo concettuale unitario, «paradiso [nelle ben note parole di Hilbert] da cui nessuno ci deve scacciare».

In questo senso, i lunghi dibattiti che nel corso del suo sviluppo il formalismo ebbe ad intrecciare con l'intuizionismo brouweriano costituiscono una nuova fase di quella lotta che già all'inizio del secolo aveva contrapposto il realismo (programmatico) di Russell al predicativismo di Poincaré e della scuola francese. Anche in questa nuova fase sono due diverse concezioni che si trovano contrapposte: da una parte l'approccio costruttivista che mira esplicitamente alla costruzione di una matematica che sia fondata non più su schemi logico-linguistici esterni, ma su contenuti dati da costruzioni mentali; dall'altra l'approccio «classico» che concepisce il problema dei fondamenti come giustificazione di una matematica già esistente (quella classica appunto) e attraverso l'analisi dei suoi procedimenti dimostrativi e definitori cerca di portarla al di là di ogni dubbio, di renderla «affidabile» ricostruendola entro un sistema con modalità «sicure». Il fatto nuovo che il formalismo porta è nel *modo* in cui questa giustificazione viene concepita, un modo che rende il confronto con il costruttivismo brouweriano più diretto e stringente.

Per il logicismo, l'abbiamo visto, la via per fondare la matematica era quella di mostrarne la natura logica e quindi garantita a priori: l'affidabilità della matematica classica diveniva quindi un corollario immediato della semplice circostanza che, essendo la logica il canone stesso della legittimità, ogni legge logica risulta di per sé legittima. La travagliata storia della Grande Logica di tradizione fregeana da Russell a Ramsey mostrava però chiaramente che i principi necessari per una ricostruzione della matematica classica trascendevano di gran lunga le pure assunzioni logiche: riducibilità, infinito, scelta, erano assiomi francamente sintetici, assunzioni la cui validità non poteva certo accettarsi come autoevidente. In questo senso la Grande Logica non offriva più garanzie di affidabilità delle stesse singole teorie particolari che voleva fondare e diveniva essa stessa una teoria matematica qualsiasi, bisognosa di giustificazione.

Del resto, già prima che con Chwistek e Ramsey il logicismo mostrasse le sue difficoltà, si erano levate voci contro il tentativo russelliano di affrontare i paradossi, via l'inglobamento di tutta intera la matematica classica e in particolare della teoria degli insiemi in un unico grande sistema logico; queste obiezioni non giungevano da parte predicativista, ma da matematici quali Zer-

melo, Hausdorff, König, che proprio richiamandosi ai lavori di Hilbert sui fondamenti della geometria invocavano una fondazione assiomatica della teoria degli insiemi, fondazione che non si ponesse come Grande Logica ma come specifica teoria matematica determinata da assiomi riguardanti oggetti matematici e nozioni logiche generali. Nello spirito, queste obiezioni non erano motivate da precise ragioni filosofiche quanto dalla netta convinzione che, essendo la teoria degli insiemi una teoria matematica e non un sistema logico, come tale andava trattata.

Vedremo in dettaglio più avanti (par. 2.2) che per Zermelo, come per König, i problemi centrali offerti dalla teoria cantoriana non erano tanto quelli delle antinomie, quanto quelli più specificamente *matematici* della giustificazione di assunzioni fondamentali come il principio di scelta. Come le obiezioni predicativiste mostravano chiaramente, il problema dei paradossi era *un aspetto* di un problema più generale: quello dell'esistenza degli enti matematici. L'«insieme» di tutti gli ordinali e l'insieme di scelta erano entrambe collezioni la cui esistenza non era accettabile dal punto di vista predicativista; solo che, se il primo era nettamente fonte di contraddizioni e matematicamente «irrilevante», non così si poteva dire per il secondo, la cui esistenza era necessaria per lo sviluppo della teoria cantoriana. È su queste basi che nel 1908 Zermelo dava una sistemazione assiomatica della teoria degli insiemi, nel tentativo di caratterizzare l'universo cantoriano imponendo assiomaticamente condizioni sulla esistenza degli insiemi: la definizione di un universo mediante specificazioni di proprietà di chiusura formulate assiomaticamente veniva così a contrapporsi alla russelliana definizione e costruzione all'interno della Grande Logica. Si trattava di una decisiva applicazione del metodo assiomatico hilbertiano che ben presto, almeno per i matematici, costituì, malgrado le lacune portate in luce da Abraham A. Fraenkel (1891-1965) e Thoralf Skolem (1887-1963) la «vera» risposta al problema delle antinomie a preferenza della soluzione russelliana.

Al fondo delle due impostazioni, quella assiomatica di Zermelo e quella russelliana, stava una dicotomia che già conosciamo dai dibattiti di Frege e Hilbert sui fondamenti della geometria e che è essenziale per capire come Hilbert nel suo programma pensasse di risolvere definitivamente il problema della fondazione della matematica. Per quanto riguarda l'affidabilità della matematica classi-

ca come pure l'esistenza degli enti matematici (due aspetti di un medesimo problema) il punto di partenza per Hilbert non è la considerazione di un sistema unitario in cui tutti i principi di formazione di oggetti matematici siano riuniti, quanto le singole teorie matematiche poste in forma assiomatica. L'affidabilità come l'esistenza matematica, in altri termini, non riguardano concetti od enti presi uno per uno ma piuttosto teorie, sistemi di enunciati.

Per Hilbert l'intera matematica non è che il repertorio dei teoremi matematici storicamente dati, repertorio logicamente articolato e diviso in rubriche opportunamente ordinate che sono le teorie assiomatiche. Non esiste contenuto specifico, intuizione fondamentale, ma una pura intelaiatura logica, una rete che collega enunciati a enunciati, teorie a teorie. Sono appunto queste ultime le unità in cui si articola il discorso matematico e non i concetti o i costrutti come per Russell e Brouwer. Scopo della matematica è fornire articolazioni concettuali per studiare il mondo, ma il mondo non è direttamente l'oggetto della matematica: occorre arricchire i dati materiali con costrutti concettuali. Fondare la matematica significa quindi fondare le singole teorie e questo a sua volta significa dimostrare la loro non contraddittorietà, la liceità cioè di assumere nello studio degli oggetti principi non direttamente giustificati nei loro termini. Le singole teorie, come sistemi di enunciati, non sono che codificazioni di condizioni che noi poniamo su enti tra loro legati da date relazioni, appartenenti al «modello» che vogliamo descrivere. In questo senso una teoria ha come unica condizione di accettabilità la noncontraddittorietà delle condizioni che pone: solo a questo patto è lecito considerare come «esistente» il dominio degli oggetti che essa postula. Esistenza degli enti matematici e affidabilità delle teorie vengono così a coincidere.

Per Hilbert, parlare dell'esistenza di un dato ente (che non sia concretamente dato) non è che usare un'espressione metaforica per dire semplicemente che le condizioni che lo specificano non sono contraddittorie: si tratta di un'esistenza ipotetica, ideale, un «come se» vincolato da una condizione materialmente significante. In questo modo i problemi tipici del dibattito tra predicativisti e logicisti si traducono in un altro problema che non riguarda più le eventuali impredicatività in cui si può incorrere nella definizione degli enti, bensì la noncontraddittorietà delle condizioni che li

determinano. Ma come stabilire questa noncontraddittorietà? È su questo punto fondamentale che Hilbert compie il passo decisivo per una risposta alle critiche intuizioniste.

La soluzione hilbertiana si basa sulla constatazione che esiste una differenza di fondo fra i vari enunciati matematici: da una parte gli enunciati la cui correttezza è immediato riconoscere, basati come sono su contenuti materiali immediatamente apprendibili, dall'altra enunciati che non godono di questa immediatezza. Si tratta di una distinzione basata sull'evidenza e verrebbe quindi facile pensare ad una sostanziale convergenza con la distinzione intuizionista fra enunciati costruttivamente significanti e non. Ma non è così; per Hilbert non esiste, come abbiamo detto, contenuto specifico – offerto da intuizioni fondamentali – proprio ad ogni singola teoria, ma piuttosto un'unica distinzione: quella fra enunciati che si riferiscono ad oggetti concreti, configurazioni spazio-temporali, ed enunciati che fanno riferimento a totalità infinite o più in generale ad oggetti astratti. Non solo, ma esiste un'ulteriore specificazione: gli enunciati immediatamente evidenti sono quelli che parlano di oggetti concreti solo per quel che riguarda loro proprietà effettivamente dominabili, concretamente verificabili mediante manipolazione. Questi enunciati fanno parte di una matematica originaria, di un nucleo primitivo del pensiero su cui non esistono dubbi, legato com'è alla semplice manipolazione concreta di oggetti, formule o diagrammi: è questa la *matematica finitista*, dotata di contenuto materiale e per la quale non si pone il problema di noncontraddittorietà.

Anche dal punto di vista intuizionista questa parte della matematica non presenta problemi. Quello che va notato è però che per gli intuizionisti la matematica «sicura», dotata di contenuto si estende al di là di questi confini, in quanto l'intuizione fondamentale della successione numerica e la nozione generale di costruzione permettono di garantire la sensatezza di teorie, quali quella dei numeri naturali (intuizionisticamente concepita) in cui il principio di induzione, alla Poincaré, risulta costruttivamente accettabile. Non così per Hilbert, che in modo più radicale pone il confine dell'evidenza immediata sul finito e non sul costruttivo. Ogni riferimento infinito non ha un immediato significato materiale e quindi richiede una giustificazione di qualche sorta: anche l'aritmetica intuizionista, al pari di quella classica

di Peano, necessita di una fondazione. I suoi enunciati riferentisi ad una totalità infinita (anche se solo potenzialmente tale) sono *ideali*, non materiali.

Tutte le teorie matematiche quindi che non fanno riferimento nei loro assiomi ai soli oggetti spazio-temporali, sono teorie ideali, prive di contenuto materiale. Per fondarle occorre dimostrare che è lecito assumere che esse abbiano un contenuto, vale a dire, in conformità a quanto detto sopra, dimostrare la loro noncontraddittorietà. Perché però questa dimostrazione abbia un significato e sia accettabile occorre che si svolga entro quella parte della matematica, la finitaria, che, avendo questo contenuto, è certa. Una dimostrazione del genere non ci porta, russellianamente, a costruire «significati» (contenuti) ma a garantire che è lecito operare «come se» questi ci fossero. Questo è un aspetto essenziale della posizione hilbertiana: il formalismo non mira a giustificare i singoli asserti o le singole teorie costruendo per essi da una base data e con metodi sicuri un contenuto, ma piuttosto vuole dimostrare che l'assunzione dell'esistenza di contenuti è materialmente giustificata su base finitista. Proprio perché l'obiettivo riguarda la teoria nel suo complesso, tutte le limitazioni motivate da considerazioni intuizioniste sulla logica che si può assumere entro le teorie perdono valore, in quanto tutto il peso della fondazione viene scaricato sui metodi che si usano nelle dimostrazioni di noncontraddittorietà, le quali ultime, proprio in quanto debbono essere giustificate in modo puramente finitista e in definitiva combinatorio, *a fortiori* soddisfano le restrizioni costruttiviste sulla logica. La differenza fondamentale è che queste restrizioni non vengono più poste sulla teoria matematica in oggetto, ma sulla «metateoria» che ha come oggetto di studio proprio la teoria.

La distinzione è profonda, in quanto porta in primo piano un concetto nuovo, quello delle teorie come oggetti e quindi della fondazione come analisi che si svolge con un linguaggio e con metodi diversi da quelli codificati dalle singole teorie. La fondazione cessa così di essere la ricostruzione entro un sistema unico per tutte le teorie e formulato nello stesso linguaggio e sullo stesso livello, per trasformarsi nello studio di esse come oggetti entro un linguaggio che non è il loro e dentro una teoria che non deve riprodurre i loro «contenuti» quanto dimostrarne la noncontraddittorietà. Il nucleo dell'approccio formalista sta appunto in questa distinzione fra un lin-

guaggio-oggetto (quello delle teorie) e un metalinguaggio (quello della matematica finitista che si configura come metateoria).

Con questo perde senso l'idea di una Grande Logica come sistemazione unitaria di tutti i concetti astratti. La distinzione fra linguaggio e metalinguaggio, fra teoria e metateoria, comporta l'abbandono del progetto di una ricostruzione globale della matematica entro una teoria che raccolga in sé tutti i principi esistenziali necessari: la matematica finitista assomma sì in se stessa il carattere fondante come la logica russelliana, ma si connette alle singole teorie non via un inglobamento, quanto piuttosto attraverso la loro assunzione come oggetto di studio. Con Hilbert si assiste quindi al frantumarsi della Grande Logica in varie sottologiche concepite questa volta come sistemi logici deduttivi da porsi alla base delle ricostruzioni formali delle varie teorie matematiche. In questo modo la nozione di logica assume quella determinatezza che Ramsey notava mancare nella sistemazione russelliana. I *Grundzüge der theoretischen Logik* (*Lineamenti di logica teoretica*), pubblicati da Hilbert e Wilhelm Ackermann (1896-1962) nel 1928, costituiscono la «canonizzazione» di questa rottura e l'avvio verso quella concezione dei sistemi logici come semplici sistemi inferenziali che è oggi corrente. In questo volume, significativamente, non viene presentato un sistema logico unico, ma diversi sistemi, quello degli enunciati, dei predicati del primo ordine, del secondo ordine ecc. ciascuno singolarmente studiato nelle sue proprietà metateoretiche. Lo stesso avviene per le teorie matematiche che vengono ora formulate entro linguaggi specifici, su diverse basi logiche, ma sempre singolarmente considerate. È una ripresa dell'approccio assiomatico dei primi anni del secolo di cui abbiamo già parlato.

Pur nelle mutate condizioni cui poi accenneremo, questa sistemazione articolata troverà la sua più completa espressione nelle *Grundlagen der Mathematik* (*Fondamenti della matematica*) di Hilbert e Paul Bernays (1888-1977) che verranno pubblicate in due volumi nel 1934 e 1939 rispettivamente, e costituiranno un punto di riferimento obbligato per l'indagine posteriore (anche relativamente ai rapporti con l'intuizionismo che solo qualche anno prima aveva cominciato a ricevere una sistemazione assiomatico-formale ad opera di Arend Heyting).

Sotto l'influenza hilbertiana, gradatamente l'interesse dei logici

si sposta, ed emergono nuovi problemi tutti legati all'idea di fondo della logica come indagine delle teorie formulate entro linguaggi formali specifici: non si tratta solo del problema della coerenza, ma anche di quello della completezza, della decidibilità, tanto delle teorie in generale quanto in particolare dei sistemi di logica pura; tutti problemi che acquistano senso solo una volta che si accetti la distinzione hilbertiana fra teoria e metateoria. Alla base di questa nuova concezione sta l'idea delle teorie come puri *sistemi formali*, oggetti combinatoriamente organizzati, appartenenti alla matematica finitista. Un sistema formale per Hilbert è dato dalla specificazione di un linguaggio, concepito come un alfabeto finito corredato da regole per la formazione dei termini e delle formule (corrispettivi formali dei nomi individuali e degli «enunciati») e da un apparato deduttivo costituito da un insieme finito, o per lo meno finitamente dominabile, di assiomi e di regole: i primi semplici formule, le seconde relazioni effettivamente specificate fra formule.

In questo modo teorie e sistemi logici trovano la loro rappresentazione materiale nei sistemi formali. Come tali essi divengono oggetti concreti, adeguati alle considerazioni della matematica finitista. È entro quest'ultima che si può formulare il problema della coerenza dei sistemi formali (delle teorie) concepito come studio delle loro proprietà combinatorie; e dal momento che la noncontraddittorietà è, fra queste, una proprietà che riguarda l'insieme delle formule dimostrabili, ne viene che il momento centrale delle considerazioni metateoriche riguarderà proprio la struttura delle dimostrazioni, e in definitiva il concetto stesso di dimostrazione proprio di ogni sistema formale. Da qui il nome di *teoria della dimostrazione* (*Beweistheorie*) che Hilbert stesso dà all'analisi, nel suo senso, dei sistemi formali. Come tale la *Beweistheorie* deve far parte della matematica finitista e ne costituisce un'applicazione: principi e metodi dimostrativi devono essere mutuati da quest'ultima che così diviene il nucleo a cui si deve ridurre, attraverso la mediazione della teoria della dimostrazione, ogni studio delle teorie matematiche.

Anche se, paradossalmente, Hilbert postulandone a priori la determinatezza non si preoccupò mai di dare una definizione esplicita di «matematica finitista», il programma generale risultava sufficientemente determinato e articolato da permettere svilup-

pi logico-matematici precisi e fruttuosi, sicché la *Beweistheorie* hilbertiana divenne per tutti gli anni venti il fulcro della ricerca logica (assieme, non va dimenticato, alla matematica intuizionista che proprio in quel periodo veniva costituendosi). Questo ruolo di guida lo perderà solo all'inizio degli anni trenta, quando la scoperta dei teoremi limitativi da parte di Kurt Gödel (1906-78) mostrerà come, nella sua formulazione originaria, il programma hilbertiano fosse destinato al fallimento: non è possibile giustificare teorie forti almeno quanto l'aritmetica senza ricorrere nella metateoria ad assunzioni più potenti, e quindi trascendenti la matematica finitista. Con questo anche il formalismo trova il suo scacco definitivo. Il risultato di Gödel ha un significato centrale nella storia della logica di questo secolo, a parte il suo enorme interesse intrinseco, proprio per questo fatto: esso costituisce la linea di demarcazione fra due epoche diverse della ricerca logica, la prima che giunge fino agli anni trenta, legata a programmi e in sostanza (se si esclude l'intuizionismo) vincolata al progetto di una giustificazione della matematica classica su basi indubitabili, la seconda che prende appunto le mosse dalla constatazione dell'impossibilità di una tale giustificazione.

Questa riconsiderazione delle alternative possibili d'altra parte non va vista come l'abbandono di programmi o l'adozione di *una pratica senza principi*; si tratta piuttosto di un approfondimento dei temi stessi alla base dei programmi e di una loro applicazione alla stessa pratica della matematica.

2. NUOVI ORIZZONTI: I PRIMI VENTI ANNI DEL NOSTRO SECOLO

Come dovrebbe risultare chiaro dal paragrafo 1, si può affermare che sostanzialmente la storia della logica nei primi vent'anni del Novecento si riflette nelle vicende delle varie scuole legate allo sviluppo dei vari programmi fondazionali (logicismo, formalismo, intuizionismo). Cominceremo la nostra esposizione con l'analisi del logicismo dopo Frege e più in particolare dei *Principia*, punto di riferimento obbligato non solo sul piano filosofico ma anche su quello tecnico. I temi che emergeranno riguarderanno la natura

di, e i rapporti fra, definizioni, dimostrazioni, costruzioni (matematiche) e in sostanza riflettono altrettanti tentativi di articolare il discorso sull'*esistenza degli enti matematici*, che resta sempre il tema centrale, direttamente e imperiosamente posto dalla scoperta delle antinomie. Come primo esito di queste discussioni si avrà in generale una separazione abbastanza netta fra quella che possiamo chiamare una concezione *logica* della fondazione, attenta alla ricerca e all'individuazione delle *assunzioni* e delle *forme di argomentazione* in base alle quali ricostruire la matematica e una concezione *matematica* della fondazione che concentra la propria attenzione sui *contenuti* della matematica, dando poi eventualmente risposte estremamente differenziate circa il modo di isolarli o di considerarli. Particolarmente interessante, in questa seconda prospettiva, la distinzione fra assunzioni e procedimenti dimostrativi. Va da sé che queste classificazioni sono in generale assai sfumate, salvo forse che nel caso degli intuizionisti.

2.1 *Il logicismo di Bertrand Russell*

La figura di Bertrand Russell è decisamente complessa e ben difficile sarebbe condensare in poche pagine gli spunti e i temi che la sua varia e copiosa produzione scientifico-filosofica offre. Qui ci dovremo limitare a entrare brevemente nel merito della sua *attività* di logico che partendo da posizioni vicine a quelle di Francis Herbert Bradley (1864-1924) lo portò come sappiamo a diventare – almeno nel primo scorcio del secolo – uno dei massimi animatori della rinata logica matematica.

Come noto, Bradley rifiutava di annettere alcuna importanza alla logica *formale* in quanto tale intendendo egli con ciò tanto la sillogistica di marca aristotelica quanto la più recente «logica equazionale» di Jevons e quindi sostanzialmente la logica matematica di Boole. Compito della logica è per Bradley quello di render conto del ragionamento in generale e da questo punto di vista, pur riconoscendo che il sillogismo era un tipo importante di ragionamento egli ne negava il ruolo esclusivo tanto più che, basandosi sulla tradizionale analisi della proposizione in soggetto/predicato, esso risultava insufficiente proprio perché la natura fondamentale del rapporto inferenziale è *relazionale*. Per ragioni

diverse Bradley rigettava la logica algebrica sostenendo che, dal momento che la logica booleana è limitata dalla sua natura matematica a trattare con problemi «che si adattano a un ragionamento numerico», essa non era in condizioni di dare alcun disegno, adeguato o no, del ragionamento in generale. La logica per Bradley non intende dare semplicemente dei *principi* del ragionamento ma effettivi canoni o tests di inferenza e d'altra parte non è possibile pensare di dare un repertorio completo delle inferenze valide, a causa del numero infinito di queste ultime. Infine – obiezione ancor più centrale – il ragionamento e l'inferenza non possono mai essere formali perché non si possono *mai* dare principi puramente formali, in quanto ogni principio contiene qualcosa di materiale. Secondo Bradley insomma «la logica formale e la matematica trattano con problemi di generalità inferiore rispetto a quella della logica come disciplina filosofica». E la causa della non accettazione della logica formale da parte di Bradley stava proprio in questa sorta di inadeguatezza filosofica che egli le attribuiva.

Nel suo primo periodo, in particolare all'epoca della pubblicazione di *An essay on the foundations of geometry* (*Saggio sui fondamenti della geometria*, 1897)[6] Russell è fortemente imbevuto di questa concezione della logica. Dichiara infatti che «in logica ho appreso per lo più da Bradley e, quindi, da Sigwart e da Bosanquet»; avendo tuttavia alle spalle una solida e profonda formazione matematica, egli voleva *dimostrare* l'essenziale irrilevanza, per la metafisica e l'epistemologia, dei moderni sviluppi matematico-formali della logica, a differenza appunto di Bradley (e di quanti altri!) che, in totale mancanza di quella preparazione, assumevano tale irrilevanza per garantita.

La posizione di partenza di Russell in questo scritto è «kantiana», nel senso che – pur con le dovute distinzioni sulle quali sarebbe qui superfluo insistere – egli vuole difendere quell'apriorità della geometria euclidea (ossia la sua necessità per l'esperienza dello spazio) che la scoperta delle geometrie non euclidee da una parte e la costituzione autonoma della geometria proiettiva dall'altra avevano come sappiamo messo in discussione. Allo scopo Russell, sulla scorta di Cayley e Klein, pone alla base dell'intera

[6] Una versione rielaborata della dissertazione di *fellowship* tenuta a Cambridge nel 1895.

costruzione alcuni principi di geometria proiettiva che egli ritiene essere a priori e che permettono di ottenere tutti i teoremi della geometria proiettiva stessa; da tali principi proiettivi egli ricava quelli metrici che «caratterizzano» tanto la geometria euclidea quanto le non euclidee.

Russell tuttavia è convinto che un altro «pericolo» alla teoria kantiana possa provenire dalla teoria del transfinito di Cantor, secondo le indicazioni di Couturat;[7] ma ancora nel 1897, quando recensisce l'opera del francese, egli ribadisce di non accettare la teoria cantoriana che, indipendentemente dalla sua utilità matematica, mai avrebbe potuto avere «validità filosofica». Tuttavia, pur ammirando gli sforzi di Couturat nel difendere «una causa così impopolare» concorda con lui su una cosa: se avesse avuto ragione Cantor allora avrebbe avuto torto Kant.

Questo, brevemente, il quadro delle posizioni del Russell «precongressuale». Avviene quindi la svolta del 1900, col congresso internazionale di filosofia e l'incontro con Peano e la sua scuola. Egli ritorna dal congresso con convinzioni filosofiche diverse e inizia l'anno stesso la stesura di *The principles of mathematics* (*I principi della matematica*) nel contempo studiando le opere di Peano, di Cantor, e rileggendo Frege, prima frettolosamente accantonato data l'inusitata difficoltà della sistemazione simbolica. Nel 1901 pubblica un articolo, *On the notion of order* (*Sul concetto di ordine*), dove sostiene che i matematici moderni, in particolare Cantor, hanno sviluppato in modo tale l'idea di ordine che i filosofi non possono più a lungo ignorare tale sistemazione, che peraltro è di centrale importanza dal momento che «i filosofi hanno in generale professato una teoria delle relazioni che, se corretta, renderebbe le successioni logicamente impossibili». Russell vuol qui rigettare il dogma filosofico fino ad allora accreditato (in particolare proprio dal suo maestro Bradley che di questo

[7] Louis Couturat, un matematico e filosofo francese (che avremo occasione di ricordare più avanti) già a quel tempo acceso sostenitore della nuova logica alla Peano e della teoria degli insiemi cantoriana, vede quest'opera di Russell come «l'Estetica trascendentale di Kant, rivista, corretta e completata, alla luce della Metageometria» e quindi decisamente ammirevole; anche se del tutto sbagliata: Couturat infatti già nel 1896 in *De l'infini mathématique* (*Dell'infinito matematico*) aveva intrapreso la refutazione della filosofia della matematica kantiana sulla base delle nuove teorie dell'infinito di Cantor.

aveva fatto un cardine della sua metafisica) che le relazioni siano in realtà stati interni delle cose, e ciò proprio in virtù del fatto che ad esempio «essere una successione indipendente è avere un posto distinto fra le entità». Di qui viene anche un deciso e ormai definitivo attacco contro la logica soggetto/predicato. La versione logica di tale distinzione non fa in effetti che riprodurre quella metafisica fra sostanza e qualità, e basta allora, per superare questo punto di vista, considerare che esiste un terzo tipo di entità: le relazioni.

Con questa nuova prospettiva Russell inizia dunque la stesura dei *Principi*,[8] ma nel 1902 scopre la famosa antinomia che comunica a Frege e che noi abbiamo visto nel capitolo precedente. L'analisi di questa antinomia lo affatica per vario tempo, fra continue oscillazioni. Comunque, quando nel 1903 il volume viene pubblicato, in un'appendice Russell propone per la sua soluzione un abbozzo di quella che sarà la teoria dei tipi. L'opera si chiude con la promessa di un secondo volume; questo non verrà mai scritto ma nel frattempo, rinsaldatasi l'amicizia e la collaborazione di Russell con Whitehead,[9] viene dai due progettata una seconda opera fondamentale, che rappresenterà il naturale punto di riferimento per almeno un ventennio di successivi studi logici. Alludiamo ai *Principia mathematica*, scritti appunto in collaborazione dai due autori, e concepiti in un primo momento come secondo volume dei *Principi*, quindi progettati in quattro volumi di cui solo tre videro la luce fra il 1910 e il 1913.

[8] Essi furono pubblicati nel 1903 come primo di due volumi, il secondo dei quali avrebbe dovuto essere «dedicato esclusivamente ai matematici» (a differenza del primo che «si rivolge in egual misura al filosofo e al matematico»). Questo secondo volume non fu mai scritto ed il suo contenuto si ritroverà nei *Principia mathematica* di cui parleremo più avanti. I *Principi* furono pubblicati in seconda edizione nel 1937 e da allora hanno conosciuto sette ristampe, l'ultima delle quali nel 1957. L'opera è divisa in sette parti più due appendici: I. Gli indefinibili della matematica; II. Numeri; III: Quantità; IV. L'ordine; V. Infinità e continuità; VI. Spazio; VII. Materia e moto. Appendice A: Le teorie aritmetiche e logiche di G. Frege; Appendice B: La teoria dei tipi. Russell afferma che l'opera ha due scopi principali. Il primo è «quello di provare che tutta la matematica tratta esclusivamente di concetti definibili in termini di un numero piccolissimo di concetti logici fondamentali e che tutte le proposizioni di tale scienza sono deducibili da un numero piccolissimo di principi fondamentali»; il secondo è «la spiegazione dei concetti fondamentali che la matematica accetta come *indefinibili*».

[9] Si veda la nota 5 a pag. 390.

La strutturazione e lo sviluppo puntuale dei volumi in questione sono oltremodo complessi e pesanti[10] e noi ce ne interesseremo solo per quanto riguarda più da vicino il nostro discorso relativo alla crisi dei fondamenti e al superamento delle antinomie. È importante però ribadire qual è l'obiettivo di fondo dell'opera: i *Principia* vogliono porsi come l'effettiva realizzazione della tesi logicista nel senso «sperimentale» di offrire una precisa ricognizione di quanta matematica si può ricostruire all'interno della Grande Logica: in altri termini, dopo che già al tempo dei *Principi* era giunto alla determinazione di accettare la teoria degli insiemi cantoriana, Russell prospetta ora con Whitehead una organizzazione logica della matematica per constatare che la sua struttura è analizzabile e ricostruibile all'interno di una teoria generale dei concetti e delle proposizioni.

È sostanzialmente questo il programma logicista di Russell e Whitehead sul cui sfondo sta appunto l'idea di una Grande Logica, una logica unificante, e unificata, che stabilito opportunamente il suo linguaggio e le sue regole di inferenza deve consentire di trascrivere e analizzare deduttivamente *tutta* la matematica. Abbiamo posto l'accento sul carattere «sperimentale» di quest'impresa non solo in riferimento al fatto che l'opera di Russell e Whitehead si presenta come la concreta esplicitazione e realizzazione di un programma, come nella vecchia situazione si presentavano i *Grundgesetze* di Frege, ma tenendo anche presente lo sviluppo successivo della logica il che comporta in certo senso un aspetto negativo dell'opera di Russell e Whitehead. È in effetti una situazione

[10] Russell dirà nella sua autobiografia che si trattò di un grossissimo sforzo intellettuale (protrattosi dal 1902 al 1910), che poté essere concluso a prezzo di molte crisi ed enorme fatica: «Così andai avanti finché il lavoro non fu terminato, ma la mia mente non si è mai riavuta del tutto dallo sforzo fatto. Da allora la mia capacità di addentrarmi in difficili astrazioni è decisamente diminuita e questo è uno, se non l'unico, motivo del nuovo orientamento delle mie attività». Alla prima edizione del 1910-13 fece seguito una seconda del 1927 con numerose ristampe, l'ultima delle quali è del 1968. I tre volumi editi (un quarto, che doveva essere dedicato alla geometria, non fu mai scritto) sono suddivisi, nella seconda edizione, come segue. Un'ottantina di pagine iniziali del I volume sono destinate alla prefazione e alle introduzioni (alla prima, rispettivamente alla seconda edizione) nelle quali vengono esposte la teoria dei tipi e quella delle descrizioni. Quindi la materia è disposta in sei parti nei tre volumi, nell'ordine seguente: I. Logica matematica; II. Prolegomena all'aritmetica cardinale; III. Aritmetica cardinale; IV. Aritmetica delle relazioni; V. Serie; VI. Quantità.

che ricorda le obiezioni di Bradley contro i logici formali: questi avrebbero stabilito dei «canoni» interpretandoli in modo tale che «chi li segue è salvo e chi non li segue è dannato». Orbene lo sperimentalismo di Russell consiste proprio in questo che, giunti a certi oggetti ultimi, siano essi proposizioni o connettivi o concetti (in particolare di tipo semantico), la logica non ha più modo di indagarne la natura o lo *status*, ma deve accettarli così come sono e costruire su di essi. Questa situazione si presenta come limitativa in particolare riguardo a tutta una serie di tentativi in direzione metamatematica che abbiamo visto nascere con Hilbert e gli studiosi di indirizzo assiomatico. Abbiamo visto come per questi studiosi esistesse uno spazio di manovra fra formulazione analitica sistematica e realtà semantica che permetteva di sviluppare indagini sui limiti e l'adeguatezza dei sistemi presentati. Per Russell e Whitehead *tutto* deve essere codificato all'interno della Grande Logica. La semantica soggiacente ai *Principia* è quella del senso comune e non vi è alcun accenno, il minimo spiraglio, che possa far pensare alla possibilità che tali concetti fondamentali possano essere ulteriormente indagati da un punto di vista logico. La granitica e complessa costruzione dei *Principia* deve comprendere tutto: al di là non si può andare. Tutto ciò che è possibile fare a questo riguardo è semplicemente porre distinzioni e rapporti, non assumere nuovi oggetti di indagine.

La necessità di evitare le antinomie è responsabile dell'enorme macchinosità della cosiddetta teoria dei tipi ramificati, che costituisce il substrato linguistico-deduttivo dei *Principia*. Con Whitehead, Russell si era deciso ad adottare tale teoria, che aveva già compiutamente esposto nel 1908, dopo vari tentennamenti che nel 1905 l'avevano portato a scartare decisamente questa soluzione prospettandone altre tre possibili nell'articolo *On some difficulties in the theory of transfinite numbers and order types* (*Su alcune difficoltà nella teoria dei numeri e dei tipi d'ordine transfiniti*) del 1905.

Si trattava: 1) della cosiddetta teoria zig-zag; 2) della teoria della limitazione di grandezza e 3) della «no-classes theory» (teoria senza classi). A noi non interessa qui presentare nei dettagli queste teorie, né discuterle in particolare, visto che come si è già detto Russell finisce poi con lo scartarle per tornare alla teoria dei tipi (che, si noti, non è neppure nominata nell'articolo del 1905). Ci limiteremo quindi a concise caratterizzazioni, per lo più dello stesso

Russell, e porremo in nota invece alcune osservazioni di Poincaré sulle stesse che ci sembra diano una sufficiente idea delle polemiche allora correnti. Come limitare il principio di comprensione fregeano?

Secondo la teoria zig-zag, dunque, «le definizioni (funzioni proposizionali) determinano una classe quando esse sono molto semplici, e cessano di farlo quando sono complicate e oscure».[11]

La teoria della limitazione di grandezza richiede invece che una classe possa esistere solo se, per così dire, «non è troppo grande»; per esprimerci metaforicamente con Russell, «non esistono cose come la classe di tutte le entità».[12] Secondo la «no-classes theory», infine, «classi e relazioni sono completamente bandite»; questo va inteso nel senso che invece di parlare di una data classe come estensione di tutti gli oggetti che soddisfano una data funzione proposizionale si può direttamente limitarsi a considerare la funzione proposizionale stessa e tutte le sostituzioni in essa.[13]

Le difficoltà inerenti a tali soluzioni (in particolare a quest'ultima, che sembrò per un certo tempo raccogliere i favori di Russell) lo portarono come già accennato a ritornare alla teoria dei tipi, che egli presentò succintamente già nell'articolo del 1908 *Mathe-*

[11] Ma chi può decidere se una definizione può essere riguardata come sufficientemente semplice per essere accettabile? Questa è la domanda che si pone Poincaré e alla quale, dice, non si trova in Russell alcuna risposta precisa o lontanamente accettabile, se non un vago appello all'assenza di contraddizione. Dunque, conclude Poincaré, la teoria è ben oscura; e, aggiunge ironicamente: «in questa notte, un solo chiarore; è la parola zig-zag». Per il concetto di funzione proposizionale o enunciativa si ricordi quanto detto nel capitolo precedente.

[12] Ancora Poincaré, esposta questa teoria, osserva che, secondo essa, «può darsi che una classe sia infinita, ma occorre che non lo sia troppo». E qui ci ritroviamo nella stessa difficoltà precedente; quando far cominciare quel «troppo»? «Ben inteso» conclude Poincaré, «questa difficoltà non viene risolta e Monsieur Russell passa alla terza teoria».

[13] In proposito è interessante riportare l'intero passo di Poincaré: «Nella *no-classes theory* è proibito pronunciare la parola *classe* rimpiazzandola con varie perifrasi. Che mutamento per i logicisti che non parlano che di classi e di classi di classi! Devono rifare tutta la logistica. Ci si figura che aspetto avrà una pagina di logistica una volta soppresse tutte le proposizioni nelle quali si parla di classi? Non resteranno che pochi sopravvissuti sparsi in mezzo a pagine bianche. *Apparent rari nantes in gurgite vasto*». Ricordiamo che il termine «logistica» venne proposto da Couturat e altri al congresso internazionale di filosofia del 1904 per designare la nuova logica matematica.

matical logic as based on the theory of types (*Logica matematica fondata sulla teoria dei tipi*), e al quale sostanzialmente ci riferiremo per schizzare una presentazione della teoria. Va premesso che all'epoca della pubblicazione di questo articolo l'elenco delle antinomie esplicitamente enunciate si era allungato di molto, dopo l'antinomia di Burali-Forti, quella di Cantor, e ovviamente l'antinomia per eccellenza, quella di Russell: si erano riesumate antiche antinomie come quella del mentitore, e altre erano state costruite. L'articolo di Russell inizia proprio con una elencazione delle antinomie fino ad allora note e crediamo convenga riprendere questo quadro per comodità del lettore. La disposizione che daremo alla presentazione delle varie antinomie segue un criterio che verrà giustificato più avanti (paragrafo 3.3).

1) Antinomia di Russell o della classe di tutte le classi che non sono elementi di se stesse. Abbiamo presentato e discusso tale antinomia nel capitolo precedente.

2) Antinomia di Burali-Forti o del massimo numero ordinale. Vale l'osservazione precedente.

3) Antinomia di Cantor o del massimo numero cardinale. Come sopra.

4) Antinomia delle relazioni. È una variante di quella di Russell e può essere formulata come segue. Sia T la relazione che sussiste fra le relazioni R e S ogniqualvolta R non è nella relazione R con S. Allora, *qualunque siano le relazioni* R e S, «R non sta nella relazione R con S» è equivalente a «R sta nella relazione T con S». Se ora diamo tanto a R quanto a S il valore T, otteniamo l'equivalenza antinomica «T non sta nella relazione T con T» è equivalente a «T sta nella relazione T con T».

5) Antinomia di Epimenide il mentitore. Si può dare in numerose forme. La più semplice ci sembra la seguente. Un uomo dice «Io mento». È chiaro che se dice la verità mente e se mente dice la verità. La proposizione in questione risulta cioè vera se e solo se è falsa.

6) Antinomia di Berry o del più piccolo numero intero non nominabile in un certo numero finito n di sillabe. Supponiamo ad esempio $n = 30$; allora l'antinomia risulta dal fatto che se si definisce «il più piccolo numero intero non nominabile in meno di trenta sillabe» si è nominato con meno di 30 sillabe un numero che non dovrebbe godere di questa proprietà.

7) Antinomia di König o del più piccolo ordinale non definibile. È stata presentata nel paragrafo 1.

8) Antinomia di Richard. È analoga alla precedente, ma preferiamo presentarla come segue (Poincaré). Consideriamo tutti i numeri decimali che si possono definire con un numero finito di parole della lingua italiana; questi numeri formano un insieme E che si vede facilmente essere numerabile (cioè di numero cardinale $\aleph_0$). Supponiamo quindi di aver enumerato gli elementi di E e definiamo un numero N come segue. Se l'n-esima cifra decimale dell'n-esimo numero di E è rispettivamente

$$0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9$$

prendiamo come n-esima cifra decimale di N

$$1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 1,\ 1,$$

rispettivamente. È allora chiaro che N non coincide con alcun elemento n di E, ossia non appartiene ad E; d'altra parte in quanto numero definito con un numero finito di parole della lingua italiana dovrebbe appartenere ad E per la stessa definizione di E.

9) Per completezza, perché verrà nominata nel seguito, illustriamo anche l'antinomia di Kurt Grelling del termine «eterologico» malgrado essa non fosse ancora nota al tempo in cui Russell scriveva il suo articolo. Supponiamo dunque di chiamare *autologico* un aggettivo (predicato) che si applica a se stesso, *eterologico* un aggettivo che non si applica a se stesso. Così ad esempio «breve» è autologico perché la *parola* (attenzione!) «breve» è effettivamente breve; «lungo» è allora evidentemente eterologico in quanto la parola «lungo» è breve, ossia non è lunga; «deutsch» è autologico (la parola è tedesca) mentre «tedesco» è eterologico; e così via. Orbene chiediamoci: l'aggettivo «eterologico» è autologico o eterologico? Se è autologico si applica a se stesso e quindi (per la definizione stessa di eterologico) è eterologico; se viceversa è eterologico non si applica a se stesso e quindi (per tale definizione) è autologico.

Nell'elenco precedente sono comprese alcune delle antinomie note nei primi dieci anni del secolo ma dovrebbe essere chiaro che esse non rappresentano che una semplice selezione da un'infinità

di antinomie costruibili. Quello che conta è che Russell ritiene di poter individuare in esse una radice comune che egli descrive come «autoriferimento o riflessività». Vediamo caso per caso. Per quanto riguarda l'antinomia di Russell, se *tutte* le classi, ammesso che non siano elementi di se stesse, sono membri di una classe *w*, lo stesso deve valere per *w*, e analogamente per la contraddizione in termini di relazioni. Nel caso del paradosso Burali-Forti (e di quello di Cantor con gli immediati e opportuni cambiamenti) la successione il cui numero ordinale causa la difficoltà è proprio la successione di *tutti* gli ordinali. L'affermazione di Epimenide deve includere se stessa nel suo proprio ambito; nel caso di nomi e definizioni, i paradossi risultano dal considerare la non nominabilità e la non definibilità come elementi nei nomi e nelle definizioni. In definitiva con le parole di Russell, «in ogni contraddizione viene detto qualcosa intorno a *tutte* le classi di un qualche tipo, e da ciò che viene detto sembra generarsi un nuovo caso che nello stesso tempo è e non è dello stesso tipo dei casi che erano *tutti* implicati in quanto viene detto».

Questa osservazione porta Russell ad escludere l'esistenza di totalità che, se ammesse come legittime, potrebbero immediatamente essere «ampliate» mediante l'aggiunta di nuovi elementi definiti in termini delle totalità stesse; sono infatti proprio totalità di questo genere che risultano presupposte, implicate in ognuno dei casi sopra esposti di antinomie. È così chiaro che Russell, riprendendolo come sappiamo da Poincaré (che l'aveva espresso per la prima volta nel 1906 nel corso di una discussione circa l'antinomia di Richard) assume come fondamentale il *principio del circolo vizioso* secondo il quale: «Tutto ciò che implica *tutto* di una collezione non deve essere un elemento della collezione» o, in modo più articolato: «Se, ammesso che una collezione abbia un totale, essa contenesse membri definibili solo in termini di quel totale, allora la detta collezione non ha totale».

Più semplicemente, non si può accettare una definizione di un ente quando in questa definizione venga implicata la totalità (la classe) cui quell'ente appartiene, o, in modo più restrittivo, quando nella definizione stessa si faccia ricorso a termini la cui definizione è possibile soltanto facendo riferimento alla classe cui l'ente da definirsi appartiene. Sarebbe ad esempio un'operazione illegittima, possibile causa di antinomie, la definizione di un numero

reale che facesse riferimento alla classe di *tutti* i numeri reali (cui ovviamente appartiene quel particolare numero che si vuol definire). Se chiamiamo *impredicative* le definizioni di questo tipo, ciò equivale in definitiva ad ammettere che l'origine di tutte le antinomie vada ricercata nell'ammissione di tali definizioni che quindi, conclude Russell, non possono essere accettate come legittime. Il nostro sistema logico dovrà allora essere congegnato in modo da non permettere la formazione di siffatte definizioni.

A questo punto è opportuno fare due osservazioni. La prima (che come si ricorderà verrà fatto in modo esplicito da Ramsey) è che le definizioni impredicative costituiscono una reale difficoltà solo per il matematico (o per il filosofo della matematica) che professi una concezione *concettualistica* della sua scienza, una concezione cioè che, qualunque sia la sfumatura con la quale viene accettata, comporti la convinzione che le definizioni (o più in generale i sistemi matematici) *costituiscano* gli enti di cui parlano. In questo caso infatti è ovvio che non è possibile riferirsi correttamente e sensatamente a una totalità che, per così dire, vado costituendo passo per passo. Nessuna difficoltà rappresentano invece le definizioni impredicative per un matematico di concezione realista: in questo caso infatti con le mie definizioni non faccio che *descrivere* oggetti che preesistono al mio atto definitorio, che sono già autonomamente costituiti e indipendenti dal linguaggio che io uso per descriverli.

La seconda osservazione è che in modo naturale, una volta accettato il principio del circolo vizioso come regolatore della possibilità di definire totalità, diviene essenziale fare una distinzione fra «tutti» (inglese: *all*) e «uno qualunque» (inglese: *any*), o se si vuole fra «tutti» inteso in senso cumulativo e «tutti» inteso in senso distributivo (ogni). È chiaro infatti che solo la prima accezione ha per così dire potere totalizzante e può condurre ad antinomie violando il principio del circolo vizioso, il che invece non avviene parlando distributivamente di un qualunque elemento in senso generico. In termini tecnici questa distinzione viene rispecchiata dall'uso del quantificatore universale (con il ricorso cioè, in termini peaniani, alle *variabili apparenti*) nel primo caso o con l'impiego delle variabili libere, non vincolate da quantificatori (*variabili reali* in terminologia peaniana) nel secondo.

Quest'ultima osservazione ci introduce direttamente nella teo-

ria dei tipi ramificata adottata da Russell e Whitehead nei *Principia*. Dal punto di vista della teoria, l'universo del discorso viene riguardato come stratificato in vari livelli o *tipi*: cosicché quando parliamo di tutti gli oggetti che soddisfano una data condizione, dobbiamo intendere solo oggetti di un dato tipo, sicché le classi di questi oggetti devono essere *omogenee* per quanto riguarda (il tipo de) i loro elementi. Allora quando ad esempio dico «tutte (*all*) le cose rosse» mi riferisco alla totalità delle cose rosse di *un certo tipo*, mentre quando dico «ogni (*any*) cosa rossa» mi riferisco *indifferentemente a una qualunque cosa rossa, indipendentemente dal tipo*: nel primo caso individuo una classe, nel secondo no. Questa tipizzazione viene effettuata per proposizioni, funzioni proposizionali, classi; in generale nel seguito ci riferiremo tuttavia alla tipizzazione in classi, che ci sembra prestarsi meglio a esemplificare intuitivamente la situazione.

L'universo viene dunque diviso in tipi nel modo seguente. Al tipo 1 apparterranno gli individui, al tipo 2 le classi di individui, ossia elementi del tipo 2 saranno *classi* i cui elementi sono individui; al tipo 3 apparterranno *classi di classi di individui*, ossia classi i cui elementi sono a loro volta classi di individui; al tipo 4 apparterranno *classi di classi di classi di individui*, ossia classi i cui elementi sono classi di classi, i cui elementi sono classi di individui; e così via per ogni tipo *finito*. Abbiamo sottolineato quel «finito» perché Russell e Whitehead affermano esplicitamente che non si possono con questa costruzione raggiungere tipi infiniti (transfiniti).

È chiaro che questa tipizzazione basta già da sola ad escludere tutta una serie di antinomie, ad esempio quella di Russell, quando si convenga che la relazione di appartenenza possa sussistere solo fra enti di tipi opportunamente diversi e si traduca questa convenzione nel linguaggio. Così ad esempio si può convenire, come fanno Russell e Whitehead, che se x^n rappresenta un oggetto di tipo n, sia sensato scrivere $x^n \in x^{n+1}$, cioè, *sintatticamente* è ben formata la formula nel caso x^{n+1} rappresenti un oggetto del tipo immediatamente successivo a n, mentre sono prive di significato (non sono cioè ben formate) espressioni quali $x^n \in x^n$ o $x^n \in x^{n+p}$ (con p intero positivo diverso da 1). La limitazione al principio di comprensione avviene così introducendo già a livello linguistico distinzioni che riflettono ambiti di significatività diversi.

Ora, nella definizione di un individuo o di una classe di un da-

to tipo può succedere di far riferimento ad altre classi (totalità), il che viene indicato, come dicevamo, dal fatto che nell'espressione definitoria compaiono variabili apparenti, vincolate cioè da un quantificatore universale o esistenziale. È proprio questa evenienza che pone le premesse per possibili definizioni impredicative; avendo accettato il principio del circolo vizioso è necessario quindi evitare situazioni di questo genere, vale a dire il linguaggio deve essere costituito in modo da escludere intrinsecamente questa possibilità. È a questo scopo che alla suddivisione in tipi si sovrappone un'ulteriore suddivisione in *ordini*, nel senso che oggetti di uno *stesso* tipo possono essere di ordini diversi a seconda delle espressioni definitorie che li individuano. Esemplificando, per comodità, per oggetti di tipo 2 (ossia per classi di individui) avremo ad esempio:

classi di tipo 2 di ordine 1	nella cui definizione non si fa riferimento alla totalità degli oggetti di *tipo* 1 (ossia nella cui espressione definitoria non intervengono quantificatori che vincolano variabili per oggetti di *tipo* 1)
classi di tipo 2 di ordine 2	nella cui definizione, oltre a eventuali variabili libere per oggetti di *tipo* 1, compare anche almeno un quantificatore che vincola variabili per oggetti di *ordine* 1.
classi di tipo 2 di ordine 3	nella cui definizione, oltre a eventuali quantificatori come sopra, compare almeno un quantificatore che vincola variabili per oggetti di *ordine* 2,

e così via per tutti gli ordini *finiti*.[14]

[14] In termini di funzioni [proposizionali] la cosa viene così espressa nei *Principia*: «Non arriveremo a funzioni di ordine infinito in quanto il numero degli argomenti e delle variabili apparenti di una funzione deve essere finito e pertanto ogni funzione deve essere di ordine finito. Dato che gli ordini delle funzioni sono definibili solo passo per passo, non vi può essere alcun processo di "passaggio al limite" e non possono darsi funzioni di ordine infinito.» La stessa cosa vale ovviamente per i tipi.

Naturalmente una tale distinzione in ordini si ha per ogni *tipo*, senza contare che la questione viene ulteriormente complicata dal fatto che ad esempio un oggetto di tipo 2 (ossia una classe di oggetti di tipo 1) può essere definita da un'espressione contenente quantificatori che vincolano variabili per oggetti di tipo 3 o di tipi superiori.

Non ci interessa proseguire nei dettagli ma è importante aver chiaro che mentre la distinzione dei tipi ci informa per così dire circa il livello dell'universo stratificato in cui si trova un dato oggetto, quella degli ordini ci dice invece «come» è stato possibile raggiungere quel dato oggetto. Per usare una felice espressione di Casari, mentre il tipo ci informa circa la complessità insiemistica di una classe, l'ordine ci informa invece circa la sua complessità concettuale. È questa sovrapposizione degli ordini ai tipi che dà la cosiddetta «ramificazione» della teoria; Russell e Whitehead fanno vedere come, tramite essa, si riesca effettivamente a evitare le antinomie, ma sottolineano anche come teoremi cruciali della matematica classica o della teoria degli insiemi non siano più suscettibili di dimostrazione nel loro nuovo contesto: per non fare che alcuni esempi, non è più possibile dare una soddisfacente definizione di identità,[15] non è più possibile dimostrare il teorema di Cantor (secondo il quale, come si ricorderà, la cardinalità dell'insieme potenza di un insieme dato è strettamente maggiore della cardinalità dell'insieme stesso), non è più possibile definire il concetto di numero reale dell'Analisi classica alla maniera di Cantor o di Dedekind, né lo stesso concetto di numero naturale alla maniera di Dedekind.

Il fatto è che, una volta «etichettato» un oggetto oltre che col tipo anche con l'ordine, non sarà più possibile parlare di *tutti* gli oggetti *tout court*, ma il quantificatore universale sarà sempre relativizzato (a un tipo e) a un ordine: in altri termini, noi potremo ad esempio avere espressioni del genere «tutte le classi *di un dato ordine* che contengono l'oggetto x» (e analogamente per il quantificatore esistenziale); mentre i teoremi o le definizioni sopra ricordate comportano un essenziale riferimento a *tutte* le classi. Nel caso

[15] Volendo infatti definire l'identità in senso leibniziano, si dovrebbe poter esprimere che due enti di uno stesso tipo sono identici quando godono delle stesse proprietà, o, in termini di classi, quando simultaneamente appartengono (o non appartengono) alle stesse classi. Ma anche qui ci si trova a dover far riferimento a *tutte* le classi, mentre il linguaggio «ramificato» ci permette di parlare soltanto di tutte le classi *di un dato ordine*.

particolare della ricostruzione dell'Analisi, come scrive Casari, la difficoltà si traduce ora «... nell'esistenza di numeri reali di ordini diversi e... nell'impossibilità in generale di trattare con sufficiente libertà gli insiemi di numeri reali. Più particolarmente, non si potrà mai parlare di tutti i reali che soddisfano una certa condizione ma sempre e soltanto di tutti i reali di un dato ordine che soddisfano quella condizione».

È proprio per superare queste difficoltà prospettate dalla ramificazione, che Russell e Whitehead introducono l'*assioma di riducibilità*, il cui senso, in modo non tecnico, può essere chiarito all'incirca come segue. Per rendere sostanzialmente inoperante la distinzione in ordini nell'ambito di un tipo, è ragionevole pensare di individuare un ordine per così dire particolarmente rappresentativo, che permetta cioè, una volta stabilita una proprietà per un certo elemento (di quel tipo e) di quell'ordine, di poterla «impunemente» estendere a *ogni* ordine. Russell e Whitehead ritengono di poter individuare questi ordini particolari in quello che per ogni tipo essi chiamano *l'ordine predicativo* di quel tipo: quell'ordine che individua enti del tipo dato facendo riferimento a elementi dell'ordine immediatamente precedente. In un certo senso ciò significa che i differenti ordini di un tipo possono essere ridotti all'ordine più basso di quel tipo. Se chiamiamo ora *predicativo* un elemento di un certo tipo (nel nostro caso, una classe) quando è determinato, descritto, da una condizione predicativa, l'assioma di riducibilità viene a dire che per ogni classe di enti di un certo tipo esiste una classe *predicativa* che è equiestensiva con essa. In termini di funzioni proposizionali in una variabile, indicando una funzione predicativa con un punto esclamativo che segue la lettera funzionale, l'assioma di riducibilità assume nei *Principia* la seguente forma:

$$\vdash \exists \mu \forall x (\phi\ x \longleftrightarrow \mu!\ x)$$

(ossia [qualunque sia la funzione ϕ] esiste una funzione predicativa μ che risulta ad essa equivalente per ogni valore della variabile x).

È chiaro allora che una volta stabilite relazioni e proprietà per classi predicative sarà possibile estendere queste relazioni e proprietà a ogni ordine grazie appunto al dettato dell'assioma di riducibilità. Ma si rivela allora anche evidente che la ramificazione viene sostanzialmente vanificata da questo assioma e a questo

punto si può innestare tutto un discorso che noi, in termini generali, abbiamo già fatto nel paragrafo 1 e che riprenderemo più avanti parlando di Ramsey e di Chwistek.

Prima di passare a trattare brevemente dell'aspetto strettamente logico dei *Principia* come assiomatizzazione della Grande Logica, è opportuno accennare brevemente agli altri due assiomi controversi che Russell e Whitehead sono costretti ad assumere. Per quanto riguarda l'assioma moltiplicativo il discorso è breve: esso afferma che ogni prodotto cartesiano di insiemi non vuoti è non vuoto e proprio in quanto gli elementi di questo prodotto non sono altro che funzioni di selezione per la data famiglia di insiemi, è immediato che esso è equivalente a quell'assioma di scelta introdotto da Zermelo nel 1904 e il cui ruolo nella matematica, come vedremo più avanti, lo stesso Zermelo avrebbe difeso mostrandone la sostanziale ineliminabilità in una concezione astratta della matematica quale si andava configurando. Dopo un iniziale scetticismo nei confronti del principio e diversi tentativi, suoi e di altri, ad esempio di P. Jourdain, Russell si convinse della necessità dell'assioma moltiplicativo per la teoria dei cardinali transfiniti, e non lo mise più in discussione.

Più sottile il discorso sull'assioma dell'infinito. La necessità di postulare esplicitamente l'esistenza di infiniti oggetti di un dato tipo, in particolare del tipo minimo, quello degli individui, era una diretta conseguenza della gerarchizzazione dell'universo russelliano. All'interno della teoria dei tipi (ramificati o no) una funzione proposizionale può avere come estensione solo oggetti di uno stesso tipo, cosicché se – seguendo la linea di Frege – identifichiamo i numeri naturali con proprietà di insiemi, quindi come classi di insiemi, avremo che perché un numero sia esemplificato, cioè non coincida con la classe vuota, è necessario che esistano nell'universo sufficienti insiemi di tipo opportuno. Russell e Whitehead identificano i naturali con la classe di quei cardinali che soddisfano il principio di induzione e hanno quindi l'onere, secondo il programma logicista, di dimostrare gli assiomi di Peano che li caratterizzano. Il problema si pone per dimostrare il secondo di tali assiomi, ossia che la funzione di successore è iniettiva. Questo assioma ha carattere esistenziale, in quanto afferma che se $m \neq n$ esisterà un insieme x di cardinalità $m + 1$ che non è di cardinalità $n + 1$. Ciò esige che si possa provare l'esistenza di insiemi di que-

sto tipo per ogni cardinale finito e quindi comporta l'assunzione che almeno un tipo sia infinito, il che – tenendo conto che esistono solo tipi finiti – comporta che infinito debba essere il tipo degli individui, il tipo «più basso».

La necessità dell'assioma dell'infinito per ricostruire all'interno della Grande Logica la matematica classica, illustra con particolare evidenza il senso in cui per Russell e Whitehead il sistema dei *Principia* deve costituire una fondazione della matematica. L'idea di fondo è che deve essere possibile tradurre ogni asserto matematico vero delle singole teorie, aritmetica, analisi, geometria, ecc. nel linguaggio dei *Principia* e qui dimostrarlo. Questo non significa beninteso che ogni singolo concetto matematico debba essere isolatamente definibile nel sistema. Come in Frege, l'unità della traduzione non è il concetto ma la proposizione. Così avremo che per non poche teorie significative noi tradurremo globalmente gli enunciati all'interno dei *Principia* ed i concetti verranno definiti contestualmente. Legata a questo processo è una delle scoperte più significative del Russell filosofo, quella teoria delle descrizioni sviluppata nel 1905 nell'articolo *On Denoting* (*Sulla denotazione*) che diventerà, con le parole di Ramsey, un «vero paradigma di analisi filosofica».

L'importanza dal punto di vista fondazionale della teoria delle descrizioni sta nel fatto che molte delle espressioni che compaiono nella pratica matematica sono – come si esprime Russell – «simboli incompleti» nel senso che non hanno una definizione assoluta, indipendente dal contesto, ma una «definizione in uso» che ne determina il significato solo relativamente a contesti specifici. Era stato Frege il primo a sottolineare la centralità di locuzioni del tipo «l'x tale che» che caratterizzano le descrizioni, e a mostrare come definizioni impossibili nella concezione classica (per genere prossimo e differenza specifica) come quella per astrazione (che permetteva di catturare la nozione di numero cardinale) portassero in modo naturale a considerare lo status logico di simili locuzioni. Il problema sorgeva per quanto riguardava l'esistenza di oggetti corrispondenti ad una data descrizione: entro che limiti possiamo affermare che esiste «la montagna d'oro» di cui parla Meinong, o «l'attuale re di Francia» di Russell o «la massima serie convergente» di Frege? In altre parole, come si analizza la logica dell'articolo determinativo? La difficoltà sta nel fatto che ci si trova di fronte a due alternative: o si assume un atteggiamento di ti-

po denotazionale, come sostanzialmente è implicito nell'uso del linguaggio naturale, per cui ognuna delle tre precedenti espressioni ha un denotato (come di fatto assumeva Meinong) oppure si è costretti a parafrasi il cui meccanismo generale è difficile da enucleare.

Frege aveva scelto la prima via introducendo un'apposita funzione la quale assumeva come valore l'oggetto denotato dalla descrizione nel caso tale oggetto esistesse (descrizione definita) un oggetto qualunque, ma determinato, in caso contrario (descrizione indefinita). La via di Russell invece è la seconda: egli introduce una procedura uniforme per tradurre contestualmente le espressioni descrittive ponendone in luce i presupposti di esistenza e unicità. Così ad esempio la proposizione «l'attuale re di Francia è calvo» va tradotta secondo Russell nella proposizione

«Esiste un unico X tale che X è re di Francia ed è calvo».

Entrambe le soluzioni hanno i loro meriti ed entrambe sono state sviluppate. Dal nostro punto di vista quello che è importante è che mediante l'analisi delle descrizioni Russell è in grado di mostrare come per specifiche teorie, quali l'aritmetica e l'Analisi, sia possibile una traduzione contestuale all'interno della Grande Logica. Un discorso diverso vale per la geometria che, come si ricorderà, diversamente che per Frege, secondo Russell è anch'essa traducibile nella Grande Logica.

Qui però il meccanismo è differente e si basa sulla concezione che opportunamente Putnam ha battezzato «*If-Thenism*» (*Se-allorismo*!) e che pone in primo piano la struttura *condizionale* implicita in quelle teorie matematiche che, come la geometria, non parlano di enti individualmente identificabili a prescindere da una struttura di fondo, ma di necessità presuppongono un contesto: i numeri, per Russell come per Frege, sono proprietà di collezioni (o di concetti), quindi enti logici singolarmente identificabili. Questo non vale per punti, triangoli, cerchi, ecc. Come aveva mostrato l'analisi assiomatica dei fondamenti della geometria di cui Russell e Whitehead raccolgono l'eredità, ogni discorso su questi oggetti ha senso solo riferendosi a determinate strutture, a determinati spazi nella fattispecie, che soddisfano certe condizioni date assiomaticamente. La possibilità di parlare di relazioni, funzioni, ecc. all'interno della Grande Logica, consente di analizzare le

strutture astratte che sottostanno al discorso geometrico e la tecnica adottata per tradurre all'interno del sistema le diverse teorie geometriche sta nell'esprimerle sotto forma condizionale: «*se* le relazioni, le funzioni, le proprietà nominate nell'enunciato geometrico $\mathscr{A}$ soddisfano le condizioni K (che si suppone costituiscano un insieme finito) date dagli assiomi che definiscono gli spazi in esame *allora* $\mathscr{A}$ deve essere vero».

Gli enunciati in questione hanno tutti la forma del tipo

$$\forall x_1 \ldots, \forall x_n (K \rightarrow \mathscr{A}).$$

È questa quella che Russell chiama una *implicazione formale*, contrapponendola ad una formula del tipo $\mathscr{A} \rightarrow \mathscr{B}$ che coinvolge solo enunciati le cui variabili non sono quantificate universalmente. Siamo giunti così ad uno degli aspetti fondamentali dal punto punto di vista logico del sistema dei *Principia*, vale a dire a quella che Russell e Whitehead chiamano la «teoria della deduzione» e che sarebbe stata il punto di riferimento per la quasi totalità delle indagini logiche fino agli anni trenta.

Già nel 1906 Russell in *The Theory of Implication* (*La teoria dell'implicazione*) aveva formulato la sua scelta di fondo per quanto riguardava le idee primitive da porsi alla base della logica e aveva optato per l'implicazione filoniana (o materiale) definita ponendo $\mathscr{A} \rightarrow \mathscr{B}$ uguale per definizione a $\neg \mathscr{A} \vee \mathscr{B}$. Tenendo conto del fatto che Russell, come sostanzialmente Frege prima di lui, assume la bivalenza, ciò significa che un condizionale $\mathscr{A} \rightarrow \mathscr{B}$ risultava falso solo nel caso in cui $\mathscr{A}$ fosse vero e $\mathscr{B}$ fosse falso. Era la stessa scelta di Frege e di fatto anche di Peirce. Nei *Principia* gli assiomi riguardanti i connettivi (quella parte della logica che Russell chiama teoria delle proposizioni elementari) ribadiscono questa opzione. I connettivi scelti come fondamentali da Russell e Whitehead sono $\neg$ e $\vee$ e gli assiomi sono:

1. $\vdash p \vee p \rightarrow p$
2. $\vdash q \rightarrow p \vee q$
3. $\vdash p \vee q \rightarrow q \vee p$
4. $\vdash p \vee (q \vee r) \rightarrow q \vee (p \vee r)$
5. $\vdash (q \rightarrow r) \rightarrow (p \vee q \rightarrow p \vee r)$.

L'unica regola assunta è il modus ponens. Non ci interessa vedere come successivamente, nel corso dei *Principia*, la teoria delle proposizioni elementari (combinazioni di proposizioni atomiche mediante connettivi) sia poi estesa utilizzando i quantificatori alla logica del primo ordine in cui si quantifica solo su individui e quindi ad una del secondo ordine (in cui si quantifica anche su funzioni proposizionali del primo ordine) ecc. per ogni tipo finito. Dal punto di vista formale va solo sottolineato come la presentazione dei *Principia* costituisca un innegabile passo indietro rispetto alla sistemazione fregeana. Quello su cui vogliamo fermarci è piuttosto il significato logico della scelta dell'implicazione materiale e del suo intergioco con quella stretta, in quanto sarà questo uno dei punti di più acceso dibattito sia dopo la pubblicazione della prima edizione (1910-'13) che della seconda (1927) dei *Principia.*

Come ricordato sopra, un'implicazione $\mathcal{A} \rightarrow \mathcal{B}$ risulta falsa solo se $\mathcal{A}$ è vera e $\mathcal{B}$ falsa. Corollario immediato è che date due qualsiasi proposizioni $\mathcal{A}$ e $\mathcal{B}$, indipendentemente dal loro significato è vera una delle due formule $\mathcal{A} \rightarrow \mathcal{B}$ o $\mathcal{B} \rightarrow \mathcal{A}$, il che risulta per lo meno singolare se leggiamo il connettivo $\rightarrow$ in termini di deducibilità, conseguenza, ecc. Come già McColl – si ricorderà dal paragrafo 5 del capitolo II – aveva obiettato a Russell, l'implicazione materiale sembra non essere in grado di rendere l'idea di una connessione valida in forza della pura forma, ossia indipendente dai valori di verità, come sembrerebbe richiesto da una nozione di implicazione logica. L'idea di McColl era quella di introdurre distinzioni modali, e quindi in un certo senso intensionali, non riconducibili cioè alla pura estensione e ai valori di verità. Sarà questa l'idea che Lewis riprenderà già nel 1912, come vedremo più avanti. Ciò che interessa osservare è che Russell e Whitehead, almeno nella prima edizione dei *Principia* non assumono un rigido atteggiamento estensionalista e ammettono nel loro universo anche funzioni di tipo intensionale come «*a* crede che», ecc. Sarà nella seconda edizione, sotto l'influenza di Wittgenstein, che essi rinunceranno alla considerazione degli aspetti intensionali, riducendo così l'universo logico ad un universo assimilato a quello dei concetti matematici nel quale «le funzioni proposizionali di cui ci si occupa in maniera specifica sono tutte funzioni di verità» in quanto «la matematica si occupa di estensioni piuttosto che di intensioni».

Rimaneva l'obiezione di McColl. La risposta di Russell si basa essenzialmente sulla distinzione fra aspetto inferenziale dell'implicazione e giustificazione semantica della medesima. Russell ammette che un legame *formale* tra $\mathscr{A}$ e $\mathscr{B}$ non può dipendere dal solo valore di verità delle due proposizioni ma deve tenere conto dei possibili contesti in cui le proposizioni assumono valori di verità. È questo il ruolo della implicazione formale di cui sopra abbiamo parlato. Asserire che $\mathscr{A}$ implica formalmente $\mathscr{B}$ significa asserire la verità di tutte le implicazioni materiali che si ottengono facendo variare le variabili sull'universo, mantenendo ovviamente le opportune distinzioni di tipo. Alle «circostanze possibili» di McColl vengono quindi sostituite le totalità dei possibili valori delle variabili coinvolte. Sullo sfondo – come avrebbe osservato Wiener nel 1912 rispondendo ad analoghe critiche di Lewis – sta l'idea che tutto il possibile ha di fatto esistenza attuale nell'universo della Grande Logica, ed in quest'ottica risulta chiaro in che senso il sistema dei *Principia* tendesse a presentarsi come onnicomprensivo e totalizzante, spiegandoci la solidarietà che venne ad instaurarsi fra i suoi oppositori, che sostenevano – per usare la terminologia di Sheffer – una visione multilocale e multibasale, secondo la quale cioè non esiste un unico universo logico e un unico stock di concetti logici fondamentali. In questa prospettiva l'implicazione materiale serviva solo per esplicitare il meccanismo inferenziale attraverso la regola del modus ponens e come tale non aveva in sé una giustificazione in termini di «legami concettuali, connessioni semantiche», ecc. Nel bene e nel male la logica dagli anni venti in poi fu con il sistema dei *Principia* che dovette misurarsi e vedremo ben presto come questo confronto abbia portato al frantumarsi dell'idea di una Grande Logica e all'emergere dei frammenti (linguaggio proposizionale, del primo e del secondo ordine, ecc.) che di fatto erano già separati all'interno del sistema ma non ne erano autonomi.

Diversa la situazione per quanto riguarda i *Principia* dal punto di vista fondazionale, cioè come formulazione di una teoria in grado di ricostruire in sé tutta intera la matematica. Come abbiamo più volte osservato, non solo i *Principia* si presentano in una veste assai più sciatta del sistema fregeano, ma la doppia gerarchizzazione in tipi e ordini era fonte di pesantezze notazionali e di ambiguità, prima fra tutte la necessità di duplica-

re nozioni come quella di identità, di numero naturale, ecc. a seconda dei tipi. Per la logica stessa che sta al fondo della nozione di tipo, ogni nozione fa riferimento ad una specifica totalità di argomenti e non si può percorrere trasversalmente tutto l'universo. Così ad esempio se parliamo del numero *m* potremo riferirci ad una proprietà che riguarda classi di un dato tipo: è indifferente quale esso sia ma non possiamo concepire totalità non omogenee. Avremo così numeri naturali in ogni tipo maggiore di quello degli individui e questa è una duplicazione che sembra completamente *ad hoc*, più legata all'impalcatura logica adottata che alla natura del concetto di numero. A questa difficoltà Russell e Whitehead sopperivano mediante l'ambiguità dei tipi, ammettendo cioè di poter parlare di tipi indeterminati. È un'idea questa che sarebbe poi stata sviluppata a partire dagli anni trenta da Quine e poi da Specker. Un altro tentativo di superare i valori imposti dalla necessità di utilizzare totalità omogenee fu fatto negli anni trenta da Carnap introducendo l'idea di tipi transfiniti, ad esempio utilizzando un tipo in cui è possibile riprodurre copie opportune di tutti i tipi finiti e quindi, attraverso queste copie, considerare totalità che non sarebbero omogenee. Questa idea si ricollega a quella gerarchia dell'universo degli insiemi che von Neumann avrebbe definito negli anni venti e che per certi versi è implicita già nell'assiomatizzazione di Zermelo.

C'è però una semplificazione ulteriore che Norbert Wiener (1894-1964) presentò nel 1914, più o meno contemporaneamente e indipendentemente da Hausdorff e Kuratowski e che permette di ricondurre la nozione di relazione a quella di insieme. Le relazioni in estensioni sono insiemi di *n*-uple ordinate e un'*n*-upla è determinata non solo dagli elementi che la compongono ma anche dal loro ordine. Ciò significa, nel contesto del sistema dei *Principia*, che la distinzione fra funzioni unarie e *n*-arie è irriducibile, cosicché – considerando i tipi come domini di significatività delle funzioni proposizionali – alla distinzione fra tipi e ordini va sovrapposta la distinzione in arietà distinte. Quello che Wiener fece vedere è che il concetto di coppia ordinata (e quindi mediante iterazione, quello di *n*-upla) è definibile in termini della pura nozione di appartenenza o se si vuole in termini di funzioni proposizionali unarie. L'idea è di considerare che due coppie $<a, b>$ e $<c, d>$ coin-

cidono se e solo se $a = c$ e $b = d$. È naturale allora porre come definizione della coppia $<a, b>$ l'insieme

$$\{\{a\}, \{a, b\}\}$$

ed è immediato verificare che se definiamo in questo modo le coppie $<a, b>$ e $<c, d>$, i due insiemi coincidono proprio nel caso in cui $a = c$ e $b = d$. È un'applicazione del metodo fregeano di definizione, la cui importanza è difficile sopravalutare nella sua semplicità. L'immagine dell'universo delle proprietà o degli insiemi ne risulta enormemente semplificata in quanto adesso – se trascuriamo la ramificazione in ordini – è solo il passaggio da un tipo alla sua potenza che ci fa passare da un livello al successivo e la relazione centrale diviene quella di appartenenza, nei termini della quale *tutti* i concetti insiemistici sono definibili, anche quello di relazione e funzione.

Esistono così formulazioni della teoria dei tipi in cui questa semplificazione viene assunta in modo che i tipi possono essere indiciati semplicemente dai numeri naturali. Questo non è completamente accettabile, come vedremo, dal punto di vista costruttivista, ma è estremamente significativo in una prospettiva realista in quanto lega in modo molto più stretto universo dei tipi ed universo degli insiemi. È appunto ai primi tentativi di formulazione assiomatica della teoria degli insiemi che ora ci volgeremo.

2.2 *Il «primo» Hilbert e Zermelo*

Abbiamo già parlato dell'introduzione di una nuova, moderna concezione dell'assiomatica da parte di David Hilbert, concretamente applicata, nelle sue *Grundlagen der Geometrie* (*Fondamenti della geometria*) del 1899, a una assiomatizzazione della geometria elementare. Come si ricorderà nel 1900, nell'articolo *Über die Zahlbegriff* (*Sul concetto di numero*), Hilbert presentava un sistema di assiomi per i numeri reali; e ancora in quell'anno, nella sua celebre relazione *Mathematische Probleme. Vortrag gehalten auf dem internationalen Mathematiker Kongress zu Paris 1900* (*Problemi matematici. Conferenza tenuta al congresso internazionale dei matematici a Parigi, 1900*) poneva fra i problemi che sfidano il mondo matematico quello della dimo-

strazione della noncontraddittorietà (o coerenza) di quel sistema di assiomi, dopo aver mostrato come anche il problema della coerenza della geometria fosse riconducibile al problema della coerenza del sistema dei numeri reali, come è fatto in dettaglio nel cap. 2 delle *Grundlagen*.

Dopo la pubblicazione dell'antinomia di Russell è chiaro che la questione della noncontraddittorietà della teoria dei naturali e di ogni sistema fondazionale acquistava dal punto di vista di Hilbert ancor maggiore e fondamentale significato sicché anch'egli interveniva nella discussione sui fondamenti con un articolo del 1904 dal titolo *Über die Grundlagen der Logik und der Arithmetik* (*Sui fondamenti della logica e dell'aritmetica*) che forma oggetto di una comunicazione presentata al terzo congresso internazionale dei matematici, tenuto ad Heidelberg dall'8 al 13 agosto 1904. È un articolo che ha un'importanza del tutto particolare, perché non solo rappresenta la prima enunciazione, dopo la scoperta delle antinomie, di un tentativo di soluzione delle stesse tramite il metodo assiomatico e una connessa dimostrazione di coerenza, ma anche perché è l'unica presa di posizione in questo periodo di Hilbert sul problema. Egli infatti tornerà sull'argomento relativo ai fondamenti della matematica solo oltre dieci anni più tardi, prima nel 1917 con *Axiomatisches Denken* (*Il pensiero assiomatico*) e quindi negli anni venti con numerosi lavori che vedremo in un prossimo paragrafo. Già in questo articolo del 1904 tuttavia vengono avanzati metodi, temi e problemi che saranno poi centrali nell'effettiva presentazione della matura *Beweistheorie* hilbertiana e soprattutto illustrano una concezione del ruolo del metodo assiomatico che non sarà solo di Hilbert ma di altri, primo fra tutti Zermelo, che proprio attraverso il metodo assiomatico cercherà di risolvere il problema delle antinomie.

Per prima cosa Hilbert chiarisce la fondamentale differenza che esiste nel caso ci si impegni a dimostrare la coerenza della geometria e quella invece dell'aritmetica. È proprio nel secondo caso, ossia relativamente all'aritmetica, che – secondo Hilbet – si hanno le diversificazioni nelle posizioni dei vari studiosi: «infatti alcune delle difficoltà nei fondamenti dell'aritmetica sono di natura diversa da quelle che si sono dovute superare quando si sono stabiliti i fondamenti della geometria. Esaminando i fondamenti della geometria ci fu possibile lasciare da parte certe difficoltà di natura

puramente aritmetica; ma non sembra lecito ricorrere ad alcun'altra disciplina fondamentale quando si tratta di affrontare la questione dei fondamenti dell'aritmetica». Le principali difficoltà in questo senso vengono messe in luce da Hilbert mediante una breve disamina critica dei diversi punti di vista.

Per cominciare, secondo Hilbert non si può affrontare il problema dal punto di vista *dogmatico* di Kronecker, che vede sì nel concetto di numero intero la reale fondazione dell'aritmetica, ma considera tuttavia i numeri naturali come dati immediatamente, eliminando così alla base la possibilità e la necessità di condurre ulteriori indagini sugli stessi;[16] né è possibile un tentativo di fondazione *empirista* come quello di Helmholtz, la cui pretesa che i numeri vengano astratti dall'esperienza viene secondo Hilbert superata dalla semplice constatazione dell'esistenza di numeri «arbitrariamente grandi», che non è possibile in alcun modo mutuare dall'esperienza. Particolarmente profonde sono per Hilbert le ricerche di Frege e Dedekind: ma il primo, accettando senza limitazioni il principio di comprensione, si espone come noto all'insorgenza di paradossi; analogamente il secondo, perseguendo un tipo di fondazione *trascendentale* (perché nella dimostrazione dell'esistenza di insiemi finiti fa ricorso alla nozione di totalità di tutti gli oggetti), si espone all'insorgere di una «inevitabile contraddizione». Anche il tentativo di Cantor di superare le antinomie a lui note mediante la distinzione fra molteplicità coerenti e incoerenti trova il suo limite nella *soggettività* dei criteri proposti per operare questa distinzione. Con questo Hilbert mostrava con estrema lucidità come l'indeterminatezza dell'universo degli insiemi testimoniata dalle antinomie ponesse in difficoltà la via tradizionale per provare la coerenza di sistemi assiomatici mediante la costruzione di modelli (l'unico metodo accettabile per Frege, come si ricorderà).

Le difficoltà messe in luce possono essere superate a parere di Hilbert ricorrendo ad una nuova concezione del metodo assiomatico. A noi non interessa qui seguire passo passo la costruzione del sistema che Hilbert presenta bensì accennare a quella che è l'es-

[16] Vedremo che nella misura in cui gli intuizionisti si rifaranno a Kronecker, articolandone tuttavia il discorso in modo sistematico, non sarà così facile per Hilbert «eliminare» la loro posizione.

senza del metodo proposto in questo articolo per dimostrarne la coerenza. L'idea di fondo è che la coerenza è una proprietà della famiglia dei teoremi di un sistema e che questa totalità può essere dominata – proprio per come i sistemi sono definiti – applicando metodi induttivi. Supponiamo allora di avere assunti certi assiomi relativi a oggetti extralogici, ed esplicitate delle regole di deduzione. Per provare che ogni teorema gode di una data proprietà *P*, si comincerà col mostrare che gli assiomi godono della proprietà e si fa quindi vedere che le regole di derivazione accettate «conservano» *P*, danno cioè luogo a teoremi che sono proposizioni che godono di quella proprietà. Per provare la coerenza basta allora mostrare che una eventuale proposizione contraddittoria *non* godrebbe di quella proprietà sicché, dal momento appunto che le regole la conservano, dai detti assiomi non può dedursi alcuna contraddizione. In questo caso possiamo dire che le nozioni definite implicitamente dagli assiomi sono coerenti o «esistono coerentemente». Il metodo presentato da Hilbert è chiaramente una sorta di dimostrazione per induzione. Questo punto sarà centrale per le critiche che Poincaré muoverà a questo primo tentativo hilbertiano: come è possibile giustificare tramite un'induzione la coerenza di un sistema assiomatico come quello dell'aritmetica, di cui proprio l'assioma di induzione è una proposizione fondamentale? Vedremo più avanti come Hilbert giustifichi dal punto di vista epistemologico questo metodo. Quello che ci preme sottolineare fin da ora è che sul piano tecnico Hilbert aveva individuato una via nuova per provare la coerenza, una via diversa da quella tradizionale basata sulla esibizione di modelli. È questa possibilità che per Hilbert pone il metodo assiomatico al centro di ogni indagine sulle proprietà (non solo sulla coerenza) delle teorie: l'assiomatizzazione è l'unica via per analizzare il problema della esistenza degli enti matematici sotto la forma di problema metamatematico di coerenza.

Elemento importante dell'articolo del 1904 è come dicevamo la presentazione, se non organica e sistematica non per questo meno esplicita, di quelle che saranno le caratteristiche distintive del metodo formalista per le dimostrazioni di coerenza, e che si possono sintetizzare in tre punti:

1) Necessità di sviluppare parallelamente logica e matematica: «L'aritmetica», ricorda Hilbert, «è spesso considerata una parte

della logica e le tradizionali nozioni logiche fondamentali sono usualmente presupposte quando si tratta di stabilire una fondazione dell'aritmetica. Se tuttavia osserviamo attentamente, ci rendiamo conto che nell'esposizione tradizionale delle leggi della logica sono già usate certe fondamentali nozioni aritmetiche, per esempio il concetto di insieme e, in certa misura, lo stesso concetto di numero. Così ci troviamo involti in un circolo vizioso e ciò accade perché se si vogliono evitare i paradossi è necessario uno sviluppo parzialmente simultaneo delle leggi della logica e di quelle dell'aritmetica.»

2) Riduzione dell'aritmetica a un sistema di formule sulle quali operare per derivazione.

3) Assunzione dell'esistenza extralogica di «oggetti fondamentali» che sono in generale «oggetti di pensiero» e delle loro combinazioni.

Hilbert ritiene di poter così superare le deficienze messe in luce nell'opera dei suoi predecessori. Punto centrale è che, una volta dato il sistema d'assiomi, si osservi che «negli assiomi gli oggetti arbitrari che prendono il posto delle nozioni "ogni" o "tutti" della logica ordinaria rappresentano solo quegli oggetti di pensiero e le loro mutue combinazioni che a questo stadio sono presi come primitivi o che debbono essere ancora definiti», il che ovviamente comporta una limitazione alla generalità fregeana fonte di antinomie. Non solo, ma «nella derivazione di conseguenze dagli assiomi gli oggetti arbitrari che occorrono negli assiomi possono quindi essere sostituiti solo con tali oggetti di pensiero e con loro combinazioni. Va inoltre notato che quando si aggiunge un oggetto del pensiero e lo si prende come primitivo, gli assiomi previamente assunti si applicano a una più ampia classe di oggetti e devono essere opportunamente modificati».

In questo modo il metodo assiomatico consente i più sottili controlli per quanto riguarda il rapporto di «conseguenza» e quindi la possibilità di assegnare, tramite la dimostrazione di noncontraddittorietà, «un significato definito e un contenuto» a nozioni («oggetti del pensiero») che di per sé ne sono prive. Nozioni di questo tipo sono ad esempio per Hilbert «insieme», «rappresentazione», «trasformazione», «relazione», «funzione», «infinito» ecc.; per tutte queste si devono considerare opportuni assiomi e col metodo precedente si possono riconoscere

essere «coerentemente esistenti». Naturalmente possiamo così dimostrare anche l'*incoerenza* di talune nozioni. In entrambi i casi – e questo è un ulteriore punto centrale dell'approccio hilbertiano – ciò che dobbiamo fare è «... considerare la dimostrazione stessa come un oggetto matematico, precisamente [come] un insieme finito i cui elementi sono connessi mediante proposizioni che affermano che la dimostrazione conduce da... [certi assiomi a un dato teorema]».

Vedremo nel prossimo paragrafo come questo studio delle dimostrazioni divenga negli anni venti lo strumento principale del programma formalista. Quello che ci preme osservare ora è come questa idea di utilizzare il metodo assiomatico per affrontare problemi connessi con l'esistenza di enti matematici, in particolare degli insiemi, si ripresenti in questo stesso giro d'anni, sotto l'influenza di Hilbert, in Ernst Zermelo.[17] L'idea di fondo di Zermelo è che il concetto centrale da chiarire non è il rapporto fra elemento e insieme, quanto quello di universo, *dominio* degli insiemi.

Già nella presentazione hilbertiana nell'articolo sopra citato, il concetto fondamentale è quello di insieme «... sicché contrariamente alla concezione attuale, la nozione di elemento di un insieme appare solo come un prodotto susseguente della nozione stessa di insieme». Il primo contributo di Zermelo alla teoria degli insiemi è un articolo del 1904, *Beweis, dass jede Menge wohlgeordnet werden kann* (*Dimostrazione che ogni insieme può essere bene ordinato*), in cui forniva una dimostrazione della proposizione che Cantor aveva creduto a più riprese di aver provato dopo averla congetturata, ossia la proposizione circa la confrontabilità di cardinali, o se si vuole

[17] Non si può collocare Zermelo in una scuola determinata delle tre che conosciamo; anzi è proprio a suo riguardo che alcuni individuano un quarto indirizzo, «insiemistico» o «cantoriano». In effetti però, anche se Zermelo agisce come matematico puro, che tenta di risolvere le difficoltà prospettate dalle antinomie con la presentazione di un sistema matematico operativo, egli condivide in un certo senso con Russell l'atteggiamento circa il contenuto del sistema matematico (e la posizione realista, anche se molto ingenua, di fondo), con lo stesso Poincaré una concezione potremmo dire kantiana della matematica come fatto globale e infine con Hilbert almeno la fiducia nel metodo assiomatico, se non addirittura la necessità dello stesso. D'altra parte poiché il grosso impatto di Zermelo nella ricerca successiva deriva essenzialmente dalla sua costruzione del primo sistema assiomatico per la teoria degli insiemi, abbiamo ritenuto lecito e opportuno associarlo a Hilbert.

del fatto che ogni cardinale è un alef, in altre parole, che ogni insieme è bene ordinabile.

Zermelo riusciva nello scopo in modo assai semplice perché assumeva per la prima volta esplicitamente un nuovo principio, che abbiamo già ripetutamente ricordato, noto come postulato (o principio) delle infinite scelte, postulato di Zermelo, assioma moltiplicativo o semplicemente come *assioma di scelta* (*AS*). Zermelo dimostra appunto che questo principio implica il teorema del buon ordinamento. Era proprio questa convinzione della necessità di *AS* per la dimostrazione di questo teorema a conferire a questo assioma la sua giustificazione, non un suo presunto riferimento intuitivo. È questo, a parte l'andamento della dimostrazione, che è per noi interessante; Zermelo, riconosciuto a Erhard Schmidt il merito di avergli suggerito l'impiego del principio di scelta, afferma: «La presente dimostrazione si basa... sul principio che anche per una totalità infinita di insiemi esistono rappresentazioni che associano a ogni insieme uno dei suoi elementi o, espresso formalmente, che il prodotto di una totalità infinita di insiemi, ognuno dei quali contenga almeno un elemento, è esso stesso diverso da zero. *Questo principio logico non può, a rigore, essere ridotto a uno più semplice, ma esso è applicato ovunque senza esitazione nella deduzione matematica*» [corsivo nostro].

Il principio in questione sollevò immediatamente una vera selva di discussioni. Basta solo pensare che la dimostrazione di Zermelo venne resa pubblica solo qualche mese dopo che König aveva presentato al congresso di matematica quella che lui considerava una dimostrazione del fatto che il continuo non poteva essere ben ordinato. Conosciuta la dimostrazione di Zermelo, König scoprì un errore nella propria e la ritirò o meglio, come abbiamo ampiamente visto nel paragrafo 1, ne ricavò un'antinomia. Non solo: già nel 1890 Peano, in uno dei suoi fondamentali lavori *Démonstration de l'intégrabilité des équations différentielles ordinaires* (*Dimostrazione dell'integrabilità delle equazioni differenziali ordinarie*), riassumendo la dimostrazione e riferendosi a un dato passaggio, aveva affermato: «Siccome non si può applicare un'infinità di volte una legge *arbitraria* con la quale a una classe *a* si fa corrispondere un individuo di questa classe, si è formata qui una legge *determinata* con la quale a ogni classe *a*, sotto opportune ipotesi, si fa corrispondere un individuo di questa classe». In altri termini Peano aveva chiara-

mente visto il principio di scelta come principio autonomo, *ma lo aveva rigettato come principio di dimostrazione*. Ancora nel 1908, nel *Formulario*, rimaneva di questo parere: «In questa dimostrazione occorre scegliere un numero infinito di volte un cubo fra più cubi, il che non è lecito se non si dà una legge generale di scelta.» Nel 1906 nell'appendice del lavoro *Super theorema de Cantor-Bernstein et additione* (*Sul teorema di Cantor-Bernstein e appendice*) esaminando il paradosso di Richard e le osservazioni in proposito di Poincaré, affermava sostanzialmente che il principio non è impiegabile perché solo l'assunzione di elementi arbitrari in numero finito è una forma di argomentazione riconducibile al sillogismo ed in generale ad un argomentare dimostrativo. Quindi se in molti casi il principio di Zermelo è riconducibile al sillogismo va bene, negli altri in cui non si sa eliminare il postulato di Zermelo «la dimostrazione non è ridotta a forme comuni di ragionamento; la dimostrazione non è valida, secondo il valore comune del vocabolo "dimostrazione"». Per le dimostrazioni in cui intervengano infinite scelte arbitrarie Peano, per così dire, sospendeva il giudizio: «La nostra opinione è indifferente. Il teorema precedente è simile al teorema di Goldbach...» ossia è una pura congettura.

A queste e ad altre numerose obiezioni Zermelo replicò, e talora duramente, nel 1908 presentando una nuova dimostrazione del teorema del buon ordinamento (sempre basata ovviamente sul principio di scelta) in *Neue Beweis für die Möglichkeit einer Wohlordnung* (*Nuova dimostrazione per la possibilità di un buon ordinamento*), dimostrazione che ricorreva questa volta a una generalizzazione del concetto di *catena* (introdotto, come si ricorderà, da Dedekind).

Fu proprio la consapevolezza che – come nel caso dell'assioma di scelta – occorreva estrarre dall'idea di insieme princìpi che determinassero la natura dell'universo insiemistico a portare Zermelo a tentare una fondazione assiomatica della teoria. Prima però di presentare il sistema assiomatico di Zermelo vale la pena di soffermarci brevemente sulle risposte ad alcune di queste obiezioni perché ci sembra mettano bene in luce l'aspetto di fondo, il senso stesso del tentativo di Zermelo.

Zermelo distingue le varie obiezioni raggruppandole in quattro tipi fondamentali, due dei quali possono essere riguardati come «metateorici» ossia sono diretti contro l'assunzione del principio

di scelta o l'impiego delle definizioni impredicative, gli altri due sono più direttamente «teorici» ossia matematici nel senso che tendono direttamente a mettere in luce eventuali errori nella stessa dimostrazione di Zermelo. Da che parte stia Zermelo ci sembra risulti chiaro quando egli osserva che «il numero relativamente alto di critiche contro la mia breve nota [quella del 1904] testimonia il fatto che, evidentemente, contro il teorema che ogni insieme può essere ben ordinato si hanno forti pregiudizi. Ma il fatto che a dispetto di un esame meticoloso... nella mia dimostrazione non fu possibile mettere in luce alcun *errore matematico* e che le obiezioni sollevate contro i miei *principi* sono mutuamente contraddittorie... mi lascia sperare che nel tempo queste resistenze saranno superate».

Ma vediamo le obiezioni al principio di scelta; abbiamo accennato alle obiezioni di Peano: sostanzialmente esse consistevano nel non riconoscere al principio di scelta la «dignità» di principio *logico* di dimostrazione ad esso rivendicato da Zermelo dal momento che «dimostrare» voleva per lui dire semplicemente ed esclusivamente «ricondurre a sillogismi» (dimostrazioni finite) e il principio in questione non sempre ammetteva tale riduzione.

Zermelo reagisce duramente a questa obiezione e mostra in che senso interagiscano nella sua concezione principi logici ed assunzioni matematiche: «Per rigettare tale principio fondamentale [il principio di scelta appunto] si dovrebbe accertare che non vale in qualche caso particolare, oppure derivare da esso una contraddizione», non fare appello all'autorità di schemi certamente privilegiati e sperimentati *ad abundantiam* ma non per questo dimostrati esaustivi.

«Anche il *Formulaire* [*Formulario*] di Peano», ribadisce Zermelo, «che è un tentativo di ridurre tutta la matematica a "sillogismi"... si basa su un certo numero di principi indimostrabili». Peano riconosce che il principio di scelta non è deducibile dai suoi, ma questo è tutto per lui: «L'idea che eventualmente il suo *Formulaire* potrebbe essere incompleto proprio in questo punto, dovrebbe, dopo tutto, suggerirsi da sola e poiché in matematica non esistono autorità infallibili, dobbiamo tener conto anche di questa necessità e non rigettarla senza un esame obiettivo.» Evidentemente Peano individua i principi su cui basa il *Formulario*

analizzando i processi di inferenza riconosciuti come corretti nel corso della storia e accetta tali principi per la loro *evidenza*. Ora Zermelo ricorda che il suo principio è stato usato inconsapevolmente da molti matematici (Dedekind, Cantor, Bernstein, Schönflies, König ecc.), e ritiene l'evidenza un criterio del tutto soggettivo; tale non è al contrario il fatto che il principio sia o no *necessario per la scienza*, il che Zermelo sostiene presentando una lista di problemi matematici non risolubili senza l'impiego del principio in questione. «Bandire dalla scienza fatti e problemi fondamentali per il solo motivo che essi non possono essere trattati per mezzo di certi principi prescritti, sarebbe analogo a proibire l'ulteriore estensione della teoria delle parallele in geometria perché si è riconosciuto indimostrabile l'assioma su cui si fonda questa teoria. In effetti i principi devono essere giudicati dal punto di vista della scienza, e non la scienza dal punto di vista di principi fissati una volta per tutte».

Altro tipo di obiezione riguardava invece la questione delle definizioni impredicative, e aveva quindi come naturale controparte Poincaré, le cui posizioni come abbiamo già visto avevano fortemente influenzato anche Russell. Oltre che per il loro valore intrinseco, l'obiezione e la susseguente discussione di Zermelo assumono qui particolare interesse perché: 1) portano Zermelo a esplicitare il suo realismo per quanto riguarda la questione dell'esistenza degli enti matematici; 2) consentono di riconoscere con chiarezza una sua concezione in certo senso kantiana (in accordo in questo con Poincaré) della matematica nella sua globalità; 3) servono a mettere in luce la netta distinzione osservata da Zermelo tra principi (fatti extramatematici di contenuto generale) e dimostrazioni (fatti strettamente matematici e come tali sottoponibili ai controlli più accurati e spietati). Dicevamo dunque che Zermelo concorda con Poincaré nel ritenere la matematica una scienza «produttiva basata sull'intuizione»; egli non obietta quindi contro la polemica che in quegli anni Poincaré stava conducendo, come abbiamo già ampiamente accennato e come vedremo ancora nelle prossime pagine, contro i logicisti; ritiene tuttavia, e questo lo riguarda direttamente, che Poincaré identifichi erroneamente la teoria degli insiemi con la logistica che egli combatte e alla quale «nega ogni diritto di esistenza». Ora finché Poincaré mantiene il discorso sul piano filosofico, poco male, afferma Zermelo: in que-

sto caso, praticamente nessun matematico ha qualcosa da obiettare. Ma «egli intraprende a combattere le dimostrazioni matematiche con le armi della logica formale, avventurandosi così su un terreno nel quale i suoi avversari sono più forti di lui». Sappiamo che per Poincaré una definizione è predicativa e logicamente ammissibile solo se essa non presuppone tutti gli oggetti che dipendono dalla nozione definita, ossia che in qualche modo possono essere determinati da essa. Orbene una definizione di quest'ultimo tipo (per esempio quella di intersezione di *tutte* le catene, che è a sua volta una catena) interviene in modo cruciale nella seconda dimostrazione di Zermelo del principio del buon ordinamento, che quindi, secondo Poincaré, non è valida. Zermelo reagisce se possibile ancor più decisamente in quanto non è il principio di scelta che si metteva così in discussione, ma la correttezza stessa della sua dimostrazione, ossia l'unico e vero argomento sul quale aveva ragione di non ammettere dubbi. A Poincaré Zermelo risponde sostanzialmente: a) con una professione di realismo che, come osservavamo parlando di Russell, vanifica l'obiezione stessa delle definizioni impredicative; b) sostenendo che dimostrazioni impredicative non sono caratteristiche della teoria degli insiemi ma si ritrovano in tutta la matematica e c) ritorcendo l'accusa di circolarità alla stessa definizione di Poincaré.

In sostanza, afferma Zermelo, «un oggetto non è creato attraverso una tale "determinazione"; piuttosto ogni oggetto può essere determinato in un'ampia varietà di modi, e queste differenti determinazioni non danno luogo a nozioni identiche, ma soltanto equivalenti, ossia nozioni aventi la stessa estensione. In effetti è proprio l'esistenza di nozioni equivalenti ciò che Poincaré sembra aver trascurato nella sua critica... Una definizione può ben basarsi su nozioni equivalenti a quella da definirsi; in effetti in ogni definizione *definiens* e *definiendum* sono nozioni equivalenti e la stretta osservanza della richiesta di Poincaré renderebbe impossibile ogni definizione e quindi tutta la scienza».[18]

L'ultimo gruppo di obiezioni che Zermelo considerava gli veniva direttamente da parte matematica e tendeva a far vedere, ad esempio, come sulla base del risultato di Zermelo si potesse rico-

[18] Si noti che in questa sua risposta a Poincaré, Zermelo concorda con le critiche che allo stesso Poincaré aveva in proposito sollevato Peano.

struire l'antinomia di Burali-Forti.[19] Giungiamo così al cuore dei criteri che ispirarono Zermelo nella sua assiomatizzazione, prima fra tutte l'idea, espressa già nell'introduzione dell'articolo del 1908, che «... non è permesso trattare l'estensione di un concetto arbitrario come un insieme e che di conseguenza l'usuale definizione di insieme è troppo ampia. Ma se nella teoria degli insiemi ci limitiamo a un certo numero di principi stabiliti come quelli che costituiscono la base della nostra dimostrazione – principi che ci permettono di formare insiemi iniziali e di derivare nuovi insiemi da quelli dati – allora tutte le contraddizioni possono essere evitate».

Questi dunque sono i criteri di fondo in base ai quali Zermelo presenta nel 1908 in *Untersuchungen über die Grundlagen der Mengenlehre, I (Ricerche sui fondamenti della teoria degli insiemi, I)*, il primo sistema assiomatico di teoria degli insiemi. Da questo punto di vista Zermelo optava per una soluzione di tipo hilbertiano e non solo perché assumeva il metodo assiomatico ma anche perché, a differenza di Russell che concepiva un linguaggio logico generale per la descrizione di tutta la matematica, Zermelo si muoveva in un tipo di linguaggio che semplicemente gli serviva alla trascrizione della teoria degli insiemi (fosse poi o meno in questa possibile, a sua volta, una trascrizione dell'intera matematica). Mentre Russell vede il discorso nell'ambito di un contesto filosofico generale, Zermelo offre una soluzione di tipo matematico, operativo, alla questione delle antinomie che avevano messo in crisi una *teoria matematica*; Russell fa ovviamente un radicale discorso di tipo logicista, Zermelo di tutta la problematica relativa ha semplicemente preso atto del fatto che occorre superare delle difficoltà precise e

[19] È interessante riportare le parole conclusive di Zermelo sull'argomento: «Ad eccezione di Poincaré, la cui critica, basata sulla logica formale – una critica che minaccerebbe l'esistenza di tutta la matematica – non ha finora avuto seguito, tutti gli oppositori possono essere divisi in due classi. Quelli che non hanno nulla da obiettare alle mie deduzioni protestano contro l'impiego di un principio generale indimostrabile, senza riflettere che tali assiomi costituiscono la base di ogni teoria matematica e che inoltre quello che io ho introdotto è anche peraltro indispensabile per l'estensione della scienza. Gli altri critici che si sono convinti di questa indispensabilità per una più profonda conoscenza della teoria degli insiemi, fondano le loro obiezioni sull'antinomia di Burali-Forti, che in effetti è *senza significato* dal mio punto di vista, perché i principi che io impiego *escludono* l'esistenza di un insieme W [di tutti gli ordinali]».

lascia addirittura del tutto indeterminata la logica soggiacente il suo sistema. Proprio questo punto, tuttavia si rivelerà una lacuna non indifferente della sua costruzione.

Il problema per Zermelo è essenzialmente quello di «chiarire» il concetto di insieme facendo sì che vengano evitati insiemi che appaiono «troppo grandi» per poter essere ammessi – pena appunto la comparsa delle antinomie – dove è chiaro che non bisogna dare a quel «troppo grandi» una carica soggettiva o di «reale» confronto di grandezze come vedremo meglio più avanti (paragrafo 3.2). Per ora «troppo grande» vuol dire solo «che comporta contraddizione». Del resto anche Russell, come sappiamo, aveva pensato a una soluzione del genere quando aveva accarezzato l'idea di evitare le antinomie con la teoria della «limitazione di grandezza»: si trattava sempre e comunque di evitare le «esistenze scomode» ed è chiaro altresì che la cosa si poteva ottenere solo limitando la portata di quel principio di comprensione secondo il quale, come si ricorderà, a ogni proprietà corrisponde un insieme, e che Russell aveva appunto dimostrato contraddittorio.

Zermelo provvede a questa limitazione assumendo al suo posto un principio di *separazione* o di *isolamento* secondo il quale è ammessa la comprensione delle proprietà solo nell'ambito di insiemi già precostituiti, nel senso che la loro esistenza o è stata postulata o può essere dimostrata nel sistema. Dovrebbe risultare ovvio che questa proposizione costituisce un reale indebolimento della primitiva portata del principio di comprensione (pur conservando forza sufficiente a realizzare una certa parte significativa della teoria di Cantor) solo a patto che nel sistema che si considera non si assuma o comunque non sia possibile dimostrare l'esistenza della classe totale: «isolare» nella classe totale equivale infatti ad assumere la comprensione generalizzata. Zermelo deve quindi escludere tale classe dal suo sistema e questo gli riesce molto felicemente dimostrando un teorema che vedremo subito dopo aver presentato il sistema assiomatico nel corso di alcune considerazioni che faremo su di esso. Poiché il sistema assiomatico di Zermelo, che indicheremo con $\mathfrak{Z}$, fornì lo spunto e la base di partenza per tutta una serie di potenziamenti e precisazioni che si avranno soprattutto negli anni venti, ci riserviamo di dar una presentazione formalizzata nel paragrafo 3.2, limitan-

doci qui a descrivere ed enunciare gli assiomi in linguaggio comune e in modo informale.

Intuitivamente Zermelo vede la sua teoria riferita a un dominio B di individui o oggetti, fra i quali vi sono degli insiemi; quegli oggetti del dominio che non sono insiemi vengono detti elementi primitivi o *Urelemente.* Come relazioni fondamentali fra gli oggetti del dominio egli assume la relazione di appartenenza $a \in b$ che ovviamente può aver luogo tra insiemi (ossia nel caso che a e b siano entrambi insiemi) o fra *Urelemente* e insiemi (ossia nel caso che a sia un elemento primitivo e b un insieme); tale relazione cioè non può aver luogo fra *Urelemente* o, più in generale, nel caso b sia un *Urelement.* In termini di questa relazione fondamentale di appartenenza si definisce la relazione $a \subseteq b$ di inclusione fra *insiemi* che può sussistere solo quando a e b siano *entrambi* insiemi. Basandosi su queste relazioni fondamentali Zermelo enuncia otto assiomi che noi riportiamo in terminologia moderna pur parafrasando direttamente l'enunciazione zermeliana.

3.1) Assioma di estensionalità

Se ogni elemento di un insieme M è un elemento di un insieme N e viceversa, allora $M = N$. Si noti che questo assioma vale *solo per insiemi* e afferma sostanzialmente che un insieme è determinato dai suoi elementi.

3.2) Assioma degli insiemi elementari

Esiste un insieme (fittizio), l'insieme vuoto $\emptyset$, che non contiene elementi. Se a è un oggetto del dominio allora esiste l'insieme $\{a\}$ il cui solo elemento è a; se a e b sono due oggetti qualunque del dominio allora esiste sempre un insieme $\{a, b\}$ che contiene come elementi solo a e b. Questo assioma assicura l'esistenza di insiemi particolarmente semplici: l'insieme vuoto appunto, l'insieme unità di ogni elemento (sia o no questo elemento un insieme del dominio) e l'insieme coppia di due elementi qualunque (siano essi o no insiemi) del dominio.

3.3) Assioma di separazione

Ogniqualvolta la funzione proposizionale $P(x)$ è definita per tutti gli elementi di un insieme M, M possiede un sottoinsieme M_P che contiene tutti e soli gli elementi x di M i quali soddisfano la proprietà $P(x)$.

Questo assioma è particolarmente delicato. Esso esprime infatti quella limitazione del principio di comprensione che abbiamo sopra chiamato principio di isolamento. Il perché della denominazione dovrebbe risultare chiaro proprio dall'enunciazione dell'assioma che infatti consente di *isolare* nell'ambito di un *dato* insieme M quegli elementi $x \in M$ che godano di una data proprietà P. Oltre alle osservazioni in generale già fatte sopra se ne impongono alcune particolari relative proprio alla formulazione di questo assioma. In esso interviene il concetto assai problematico di «proprietà definita» che non è specificato da Zermelo. È questa una lacuna fondamentale della sua presentazione, che mostra l'assenza di una base logico-linguistica chiaramente ed esplicitamente precisata. Zermelo si limita a dire che «una questione o asserzione P è detta essere *definita* se le relazioni fondamentali del dominio, per mezzo degli assiomi e delle *leggi universalmente valide della logica* [corsivo nostro] determinano senza arbitrarietà se essa vale o non vale». Qui non solo risulta la necessità di un riferimento a una base logica esplicitata (quali sono queste leggi universali, come entrano nella determinazione della «definitezza», ecc.) che manca nel sistema zermeliano, ma si prospetta il problema, una volta superata questa lacuna, di vedere come definire opportunamente questa nozione. Vedremo che questo sarà uno dei problemi che più affaticherà negli anni venti gli insiemisti interessati a una fondazione assiomatica.

Già sulla base di questi assiomi, Zermelo riusciva di dimostrare il teorema che esclude la possibilità di antinomie come quella di Russell, teorema secondo il quale «ogni insieme M possiede un sottoinsieme M_0 che non è un elemento di M», vale a dire non ogni elemento del dominio può essere elemento di uno stesso insieme, e pertanto il dominio B stesso (il quale per sua stessa definizione contiene tutti gli elementi) *non* è un insieme e non può quindi applicarsi l'isolamento nella classe totale, ossia il principio di separazione è *effettivamente* più debole del principio generalizzato di comprensione, e non permette la derivazione dell'antinomia di Russell.

È soprattutto a questo terzo assioma dunque che è affidato il compito di evitare le antinomie e il passo in cui Zermelo fa quest'affermazione è interessante perché allude chiaramente a una sorta di suddivisione delle stesse, che si rivelerà fondamentale, co-

me avremo occasione di vedere nel paragrafo 3.3. «Dando un'ampia libertà nel definire nuovi insiemi – egli scrive – l'assioma [3.3] in un certo senso fornisce un sostituto per la definizione generale di insieme... Esso differisce da tale definizione poiché contiene le seguenti restrizioni. In primo luogo un insieme non può mai essere *definito indipendentemente*... ma deve essere sempre *separato* come sottoinsieme da insiemi già dati; così nozioni contraddittorie come "l'insieme di tutti gli insiemi" o "l'insieme di tutti i numeri ordinali", e con essi i "paradossi ultrafiniti", per usare un'espressione di Hessenberg, sono esclusi. In secondo luogo,... il criterio definitorio deve essere sempre definito... (ossia per ogni singolo elemento x di M le relazioni fondamentali del dominio devono determinare se esso vale o no) col risultato che, dal nostro punto di vista, tutti i criteri come "definibile con un numero finito di parole" e quindi "l'antinomia di Richard" e "il paradosso della denotazione finita" svaniscono. Ne segue anche però che, a volere essere rigorosi, prima di ogni applicazione dell'assioma III noi dobbiamo dimostrare che il criterio $P(x)$ in questione è definito; nelle considerazioni che seguono questo verrà in effetti dimostrato tutte le volte che non sarà evidente».

3.4) Assioma dell'insieme potenza

A ogni insieme T corrisponde un altro insieme $\mathscr{P}(T)$ detto insieme potenza di T, i cui elementi sono tutti e soli i sottoinsiemi di T, compreso T stesso e l'insieme vuoto.

Si può comprendere la portata di questo assioma, in particolare in riferimento alla teoria dei numeri cardinali, se si ricorda il teorema di Cantor per cui la cardinalità di T è *strettamente* minore di quella di $\mathscr{P}(T)$.

3.5) Assioma della riunione

A ogni insieme T corrisponde un insieme $\bigcup T$ che contiene come elementi tutti e soli gli elementi degli elementi di T.

3.6) Assioma della scelta

Se T è un insieme i cui elementi sono tutti insiemi mutuamente disgiunti e non vuoti, allora la riunione $\bigcup T$ include almeno un sottoinsieme S che ha uno e un solo elemento in comune con ogni elemento di T.

3.7) Assioma dell'infinito
Esiste nel dominio almeno un insieme Z che contiene l'insieme vuoto $\emptyset$ *come elemento e che con ogni suo elemento a contiene anche* $\{a\}$, *insieme unità di a.*

Zermelo è quindi il primo, ancor prima di Russell, ad avvertire la necessità della postulazione dell'esistenza di un insieme infinito. Si noti che la formulazione precedente comporta (unito all'assioma di separazione) che nel dominio B sia compreso almeno l'insieme $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \{\{\{\{\emptyset\}\}\}\}, ...\}$ che costituisce una copia isomorfa della struttura dei numeri naturali.

Per concludere la descrizione del sistema di Zermelo riportiamo le sue stesse parole sulle questioni più generali relative a tale sistema: «... intendo mostrare come l'intera teoria creata da Cantor e Dedekind può essere ridotta a poche definizioni e a sette principi, o assiomi, che sembrano essere *mutuamente indipendenti* [corsivo nostro]... Non sono ancora stato in grado di dimostrare rigorosamente che i miei assiomi *sono coerenti, malgrado ciò sia certamente essenziale* [corsivo nostro]; mi sono dovuto limitare a mettere ancora in evidenza che le antinomie finora scoperte svaniscono tutte se si assumono come base i principi qui proposti...». Sarà questo uno dei problemi che la *Beweistheorie* di Hilbert cercherà di affontare.

2.3 *I «semiintuizionisti» francesi e il «primo» Brouwer*

Nel primo decennio del nostro secolo in Francia la discussione sui fondamenti della matematica fu molto vivace e vide impegnati fra gli altri i componenti della scuola degli analisti di Parigi (i «semiintuizionisti» francesi) oltre che Poincaré, Pierre Boutroux (1880-1922), Louis Couturat, ecc. Gli scambi di idee estremamente vivaci, talora ironici, vertevano sostanzialmente su due argomenti fondamentali: la nuova logica di Peano e Russell e la teoria degli insiemi cantoriana, in particolare la questione della scelta sollevata da Zermelo.

Lo spunto alle discussioni venne senz'altro dal matematico e filosofo Louis Couturat, che si presentava un po' come l'entusiasta assertore della nuova logica, soprattutto in funzione antikantiana. La polemica Couturat-Poincaré-Russell si svolse soprattutto sulla *Revue de métaphisique et de morale*. Iniziò Couturat con la *Logique ma-*

thématique de M. Peano (*Logica matematica di Peano*, 1899), seguito nel 1904 da un lungo articolo, *La philosophie de la mathématique de Kant* (*La filosofia kantiana della matematica*) e nel 1906 da *Pour la logistique. (Réponse à M. Poincaré)* (*Per la logistica. Risposta a Poincaré*). Poincaré da parte sua vi scriveva tra il 1905 e il 1906 una serie di articoli che venivano poi ristampati con minime variazioni nei volumi *La valeur de la science* (*Il valore della scienza*, 1905), *La science et l'hypothèse* (*La scienza e l'ipotesi*, 1906) e *Science et méthode* (*Scienza e metodo*, 1908). Nella polemica, Russell intervenne nel 1906 pubblicando sempre sulla stessa rivista l'articolo *Les paradoxes de la logique* (*I paradossi della logica*). Parallelamente, in ambito più strettamente matematico, si situano le *Cinq lettres sur la théorie des ensembles* (*Cinque lettere sulla teoria degli insiemi*), uscite nel 1905 nelle quali Hadamard-Borel, Baire-Hadamard, Lebesgue-Borel, Hadamard-Borel e Borel-Hadamard si scambiavano appunto i loro punti di vista sulla teoria degli insiemi, e in particolare sull'assioma di scelta.

Abbiamo già parlato di Russell e del suo rapporto con Poincaré; di Couturat non resta[20] che dire che talora il suo entusiasmo per la nuova logica e teoria degli insiemi, professato in modo un po' trionfalistico, può aver messo in secondo piano quella che poteva essere l'elaborazione di un pensiero personale. Caratteristico comunque della sua attività filosofica è il tentativo di ribaltare, con questi nuovi strumenti concettuali, l'impostazione della filosofia della matematica kantiana, col corollario che l'intuizione sarebbe stata definitivamente bandita dalla matematica. Ma certamente alcune enfatiche e categoriche affermazioni di Couturat, che si dichiara modestamente semplice «relatore di sì grandi maestri» (intendendo Peano e Russell) non solo avrebbero probabilmente messo «in difficoltà i maestri stessi» (Poincaré) ma oltre a offrire lo spunto all'intervento di Poincaré nella polemica non gli risparmiano frecciate velenose e sarcastiche (anche se certamente del tutto fuori luogo e «antipatiche») da parte di quest'ultimo.[21]

[20] In questa sede. Altrimenti sarebbe forse non privo di interesse tentare di portare alla luce i tratti salienti del pensiero autonomo di Couturat, al di là delle sue stesse frequenti dichiarazioni di «dipendenza».

[21] A proposito dell'articolo su Kant: «Si vede bene che è il centenario della *morte* di Kant»; a proposito delle «ingenue illusioni» di Couturat sulla logistica che avrebbe messo «i trampoli e le ali» dopo che Peano aveva pubblicato, dieci anni

Ritorniamo invece brevemente a Poincaré, che rientra largamente nell'ambito della scuola degli analisti francesi, anche se con forti differenze (che comunque anche gli altri esponenti presentano fra loro), per poi vedere come già nel 1907 Brouwer interviene nella discussione sui fondamenti della matematica in modo del tutto autonomo e originale.

Se, con Heyting, consideriamo «intuizionisti» in matematica quegli studiosi che accettano i due seguenti principi: 1) la matematica non ha solo un significato formale, ma anche un contenuto e 2) gli oggetti matematici sono afferrati immediatamente dallo spirito pensante, sicché la conoscenza matematica è indipendente dall'esperienza; se accettiamo questa caratterizzazione, allora si può a buon diritto affermare che, almeno in senso lato, la scuola francese prelude a quello che oggi è noto appunto come intuizionismo. Vi sono tuttavia notevoli differenze che soprattutto riguardano la concezione dell'esistenza degli enti matematici. Da una parte infatti si può ritenere che gli oggetti matematici esistano indipendentemente dal nostro pensiero, ma che noi concludiamo alla loro esistenza soltanto per mezzo di una costruzione, o comunque di un criterio effettivo di individuazione. Questo atteggiamento è grosso modo comune ai semiintuizionisti francesi[22] (come lo era stato di Kronecker, e lo sarà di Skolem); ma già Poincaré accettava l'equivalenza esistenza = dimostrazione di coerenza. Più radicale è invece l'altra via che prenderà Brouwer, per il quale gli oggetti matematici non hanno alcuna esistenza indipendente dal pensiero umano, o comunque possono essere assunti come individui solo in termini giustificabili sul piano costruttivo.

Il problema di fondo che Poincaré affronta è sostanzialmente il seguente: è possibile ridurre la matematica alla logica senza fare appello a principi che sono propri della matematica? La scuola logicista, che si sforza di dimostrare questo con l'aiuto della mate-

prima, il *Formulario*: «Come, sono dieci anni che avete le ali, e non avete ancora volato!»; senza contare apprezzamenti personali assai più pesanti che non mette veramente conto di riportare.

[22] Con forti sfumature diverse. Secondo Lebesgue, ad esempio, per dimostrare che un insieme non è vuoto si deve *nominare* esplicitamente almeno un suo elemento. È chiaro comunque quale doveva essere, con questi presupposti, l'atteggiamento nei riguardi della stragrande maggioranza dei risultati della teoria cantoriana degli insiemi, e in particolare nei riguardi dell'assioma di scelta.

matica dell'infinito *attuale* di Cantor, è andata incontro secondo Poincaré solo a degli insuccessi, confermati dalle antinomie, e non miglior sorte ha avuto il tentativo del 1904 di Hilbert.

La ragione di questi insuccessi, anzi la ragione di principio per cui non sono possibili né la riduzione logicista né quella formalista della matematica, è sostanzialmente unica: la matematica ha alla sua base, kantianamente, giudizi sintetici a priori, che non ammettono quindi una eliminazione dell'intuizione o del «contenuto» della matematica stessa. In particolare, e questo spiega intanto l'impossibilità logicista della riduzione, il principio veramente «creativo» della matematica, quello che ci fa ottenere «più» di quanto non sia contenuto nelle premesse dei nostri ragionamenti è proprio quel principio di induzione matematica che i logicisti hanno assunto come definitorio dei numeri naturali (vedi assiomi di Peano) immettendosi così in un ineliminabile circolo vizioso. Più sottile è la critica a Hilbert. Poincaré riconosce i successi di Hilbert nell'assiomatizzazione della geometria, perché anch'egli ammette la definizione per postulati, *ma solo nel caso in cui si sappia dimostrare la coerenza dei postulati stessi*: ora ciò risulta possibile per i sistemi geometrici proprio perché la dimostrazione ha qui quel carattere relativo di cui abbiamo più volte parlato, che «scarica» sull'algebra e in definitiva sull'aritmetica tutte le difficoltà. Ma di fronte a un sistema di assiomi per l'aritmetica, dal momento che non può procedersi ad una verifica diretta a causa dell'infinità dei numeri naturali, è ancora necessario servirsi del ragionamento per ricorrenza, ossia in ultima analisi ricorrere all'induzione matematica, che è ancora una delle proposizioni che noi ci ripromettiamo di giustificare. Anche in questo caso siamo quindi di fronte a uno scacco di principio, a un evidente circolo vizioso.[23]

[23] È qui opportuna qualche considerazione che potrà essere compresa compiutamente quando parleremo del programma «maturo» di Hilbert, ma che tuttavia può essere intesa – crediamo – almeno nelle sue linee generali, tenuto conto di quanto detto nel paragrafo 1. Anche se nel 1905, quando Poincaré avanzava queste osservazioni, Hilbert non aveva ancora espresso chiaramente la distinzione fra matematica e metamatematica, Poincaré riconosce per così dire la presenza di *due* applicazioni dell'induzione matematica, una nel *sistema*, l'altra nella *metateoria* del sistema stesso; e ritiene che una dimostrazione di coerenza possa giustificare la prima di queste due induzioni, ma non certo la seconda. Hilbert risponderе a questa osservazione solo nel 1922 in *Neubegründung der Mathematik (Erste Mitteilung)* (*Nuova fondazione della matematica. Prima comunicazione*) sostenendo che solo una delle due (la seconda) è una forma di induzione irriducibile. Successiva-

D'altra parte anche la stessa giustificazione logicista dell'induzione – come si è visto – presentava nella sostanza, a livello di definizioni, quello stesso circolo vizioso che Poincaré ritrovava a livello deduttivo nell'approccio formalista. I motivi, o meglio, *il* motivo essenziale che porta a queste *impasses* viene illustrato da Poincaré a partire dal paradosso di Richard che noi già conosciamo. Lo stesso Richard aveva proposto una soluzione al riguardo; Poincaré la elabora e la fa propria presentandola come la «vera soluzione» (contro la quale, abbiamo visto, polemizzavano anche Peano e Zermelo). Riferendoci all'esposizione già data nel paragrafo 1, la risposta di Richard e Poincaré è semplice: E è l'insieme di tutti i numeri che si possono definire con un numero finito di parole e *senza introdurre la nozione dell'insieme E stesso*. Se non si osserva questa precauzione la definizione di E contiene un *circolo vizioso* dal momento che non si può definire l'insieme E tramite se stesso. Orbene, per evitare circoli viziosi di questo tipo occorre servirsi solo di definizioni *predicative*, considerando impredicative quelle definizioni che appunto contengono un circolo vizioso.

Abbiamo già visto come Russell assuma questo principio del circolo vizioso come elemento di salvaguardia delle antinomie e come invece tanto Peano che Zermelo controbattano alla pretesa di Poincaré di limitarsi, per così dire, a una matematica predicativa. Comunque, una volta stabilito questo, Poincaré conclude che è proprio l'affermata esistenza dell'*infinito attuale* che dà origine alla possibilità stessa di definizioni impredicative; è dunque

mente distinguerà fra una induzione «contenutistica» (e quindi giustificata, che è alla base della dimostrazione di coerenza) e una «formale» (quella dell'aritmetica), distinzione che tuttavia abbandonerà dopo qualche anno. La risposta di Hilbert, in quanto si limita a porre la distinzione senza determinarne le modalità, è chiaramente non soddisfacente come del resto venne fatto notare da molti autori a lui legati, ad esempio Hermann Weyl, Thoralf Skolem (che propendono per Poincaré), e Wilhelm Ackermann (che pone l'induzione fra i principi più delicati e questionabili dell'impostazione hilbertiana). Jacques Herbrand identificherà l'induzione contenutisticamente giustificata e quindi applicabile nella matematica come quella che si applica alle sole formule aperte (senza quantificatori). Il problema della induzione rimarrà un punto basilare per il formalismo, in quanto coinvolge il concetto stesso di finitismo e la distinzione tra ciò che riguarda manipolazioni concrete e ciò che rimanda all'infinito e a enti ideali astratti. In altri termini, è proprio dal tipo di induzione che viene ammesso a livello metamatematico che si può determinare la nozione di «finito» che si assume.

all'infinito attuale che in ultima analisi vanno riportate queste recenti difficoltà apparse nella logica e nella matematica: «*non esiste infinito attuale*; i cantoriani l'hanno dimenticato e sono caduti in contraddizione». Affermazione questa per lo meno singolare se si pensa che l'antinomia di Russell mostrava *ad abundantiam* come non l'infinito, ma l'autoriferimento fosse alla base dei paradossi.

Nella discussione fra Poincaré e Russell, Brouwer intervenne con la pubblicazione, avvenuta nel 1907, della sua tesi dal titolo *Sui fondamenti della matematica*,[24] buona parte della quale è appunto dedicata all'oggetto di quella disputa. Brouwer non concorda con nessuna delle parti in causa. Non è d'accordo col tipo di intuizionismo professato da Poincaré perché ritiene che la fonte reale delle difficoltà della logica (e della teoria degli insiemi, e in definitiva, come vedremo, della posizione di Hilbert) vada ricercata semplicemente, ma più radicalmente, nella «confusione fra atto di costruzione matematica e linguaggio della matematica». In altri termini, il logicista non fa che creare solo una struttura linguistica, dà origine a un complesso di *parole* che in nessun caso possono essere effettivamente calate nella matematica reale che è rigidamente *alinguistica*. E lo stesso Poincaré mostra di cadere in questo errore quando accetta l'identificazione di esistenza matematica con assenza di contraddizione. In realtà, per Brouwer, esistere matematicamente significa essere *costruito intuitivamente* e «se un linguaggio ausiliario è esente da contraddizioni non solo non ha in sé alcuna importanza, ma neppure è certamente un criterio di esistenza matematica».

I rapporti tra matematica e logica risultano ancora più chiari nella descrizione che Brouwer dà dei passi successivi che potano alla edificazione di matematica e logica. Possiamo riassumerli come segue:

Primo stadio: pura costruzione di sistemi matematici intuitivi «che se applicati sono espressi esternamente nella vita e visti nel mondo come matematici»;

[24] Come ormai consueto in quest'opera, daremo nella lingua originale (con traduzione relativa) i titoli di lavori in lingua tedesca, francese o inglese. Per tutte le altre (danese, russo, olandese, ecc.) riporteremo il titolo direttamente in italiano.

Secondo stadio: aspetto *linguistico* della matematica, ossia il linguaggio, scritto o parlato, della matematica;

Terzo stadio: in esso lo stesso linguaggio diviene oggetto di studio, ossia si ha una visione *matematica* del linguaggio, del quale si notano le costruzioni e strutture logiche, fermo restando tuttavia il fatto che le costruzioni di relazioni fra questi elementi sono ancora puramente matematiche;

Quarto stadio: è lo stadio nel quale, non considerando il significato degli elementi linguistici dello stadio precedente, si può costituire una «matematica del secondo ordine» che è appunto un sistema logistico;

Quinto stadio: è per così dire lo stadio metateorico del precedente nel quale cioè viene formulato esplicitamente un linguaggio in cui esprimere i principi di un sistema logistico;

Sesto stadio: è quello che impareremo a conoscere come «metamatematica» hilbertiana, nel quale cioè la matematica stessa viene riguardata come un linguaggio, sicché si potranno considerare gli elementi di questo stadio in un

Settimo stadio: nel quale si avrà per così dire una «matematica del terzo ordine», il che porterà naturalmente al momento linguistico di questa nuova matematica, in un

Ottavo stadio: e così via all'infinito.

L'unico vero e possibile fondamento di tutta questa costruzione resta per Brouwer la *matematica «reale»* che si fonda sull'intuizione, anzi, come Brouwer dice, sulla «sovraintuizione» di uno scorrere continuo del tempo, che è il fondamento stesso del continuo misurabile del matematico. Questa intuizione fondamentale è l'intuizione «dell'unità nella differenza, della persistenza del mutamento» e del *continuo come un tutto.* Le due intuizioni sono distinte e

Brouwer ribadisce che «una costruzione [del continuo come un tutto] che crei dalla intuizione matematica tutti i suoi punti individualmente è inconcepibile e impossibile». Dopo queste prime prese di posizione in cui sembra ammettere due fonti intuitive distinte per il continuo spaziale e per la successione dei naturali, Brouwer giungerà più tardi a una completa riduzione della geometria a una base analitica, affermando che i principi della geometria non sono «sintetici a priori», dal momento che a suo parere la mente umana potrebbe applicare all'esperienza qualunque tipo di geometria essa scegliesse. Gli unici veri principi a priori e sintetici sono collegati alla «sovraintuizione» di cui si diceva prima dell'unità nella pluralità nel tempo. Base della matematica diventa allora il solo scorrere del tempo che determina in certo senso quali sono i giudizi sintetici a priori che rendono possibile la matematica stessa: la possibilità di sintesi matematica, il pensiero dell'unità nella molteplicità e la costruzione successiva indefinita di nuove unità nella molteplicità; la possibilità di «interpolazione», che si manifesta ad esempio nella concezione della connessione di due diversi elementi di uno stesso tutto; ed infine il principio di induzione matematica.

Vedremo in seguito come Brouwer ingloberà tutte queste concezioni nell'edificio di una vera e propria nuova matematica; fin da ora però è chiaro che sono le leggi della logica a dipendere dalla matematica e non viceversa; la matematica reale è un «costruire privo di parole», mentre la logica, per quanto aspetto matematico essa si dia, studia solo *parole*, non costruzioni matematiche. In questo processo di costruzione la contraddizione è impossibile; se i logicisti hanno incontrato antinomie, esse sono state originate dall'aver ritenuto valide in generale leggi logiche che in effetti possono valere solo per il *finito*, ma in quanto leggi logiche applicabili solo alle parole, non lo sono alle costruzioni matematiche in generale. Vedremo più avanti quali leggi generali Brouwer riterrà di dover mantenere: quella che mette subito in crisi è la legge del terzo escluso, che non può essere applicata al processo di costruzione matematica in quanto comporterebbe di fronte a ogni enunciato $\mathscr{A}$, la possibilità di trovare una dimostrazione di $\mathscr{A}$ oppure di $\neg \mathscr{A}$. Avremo modo di vedere quale sia la portata di questo rigetto per la matematica. Nel contesto del problema dei paradossi è comunque subito chiaro che non accettare questo principio esclu-

de le antinomie (o, per essere esatti, un certo tipo di antinomie, almeno in prima approssimazione). Limitiamoci a esemplificare la cosa per l'antinomia di Russell e per comodità ricorriamo a una presentazione simbolica. Indichiamo con R (da Russell) la classe di tutte le classi che non si appartengono, ossia $R = \{x \mid x \notin x\}$. Abbiamo allora

$$x \in R \longleftrightarrow x \notin x$$

donde sostituendo R a x

$$R \in R \longleftrightarrow R \notin R \qquad (1)$$

Per la legge del terzo escluso però

$$R \in R \vee R \notin R. \qquad (2)$$

Ora se $R \in R$, allora $R \notin R$ per la (1); se $R \notin R$ allora, sempre per la (1) $R \in R$. In entrambi i casi $R \in R$ e $R \notin R$ e quindi per la (2)

$$R \in R \wedge R \notin R$$

ossia la contraddizione.

La spiegazione dell'insorgere delle antinomie che Brouwer fornisce consiste allora semplicemente nel sottolineare il ruolo della cruciale legge (2), mostrando così come il paradosso sia una semplice trappola dovuta a convenzioni puramente linguistiche. Limitiamoci per ora a queste prime considerazioni sull'intuizionismo brouweriano nel suo carattere più specificamente «distruttivo»; della *pars construens* della logica e della matematica neointuizionista parleremo più avanti nel contesto della ricerca negli anni venti.

2.4 *Riepilogo. La scuola italiana*

Possiamo gettare ora uno sguardo riassuntivo su questo primo, cruciale periodo della costituzione della nuova logica. Sono ormai emersi gli indirizzi che caratterizzeranno le ricerche almeno fino agli anni trenta, in particolare fino alle scoperte di Gödel. Le varie scuole hanno ormai avanzato più o meno efficacemente e compiutamente i punti centrali delle rispettive concezio-

ni e sarà storia dei prossimi anni vedere come queste saranno approfondite o comunque elaborate. In vista di ciò è forse opportuno fermarci sulle differenze tra i vari indirizzi per quanto riguarda il problema nodale dell'esistenza degli enti matematici conducendo il discorso da un punto di vista «astratto» relativamente ai *presupposti* e senza considerare le eventuali differenze fra esponenti di uno stesso indirizzo.

Innanzitutto un'osservazione preliminare relativa al rapporto tra matematica e logica già chiaramente delineatosi nelle varie concezioni. A questo riguardo si giunge, da parte dei logicisti a degli intuizionisti, a posizioni del tutto contrapposte. Per i logicisti questo rapporto è di tipo prioritario a vantaggio della logica: la matematica è riconducibile alla logica, la quale ha, nei riguardi della prima, un ruolo garante di fondazione: *tutta* la matematica è risolvibile, senza residui o principi propri, in principi e strutture logiche. Vi è quindi la mediazione, per così dire, del «formalismo» hilbertiano, per i quali è necessario non parlare di dipendenza dell'una scienza dall'altra, ma piuttosto si tratta di sviluppare parallelamente le due scienze in questione. Se infatti non è possibile eliminare dai fondamenti della logica riferimenti di tipo aritmetico, non è neppure possibile sviluppare tale disciplina al di fuori di una precisa ed esplicita sistemazione della logica nei suoi principi e nelle sue strutture inferenziali. Infine vi è il ribaltamento, rispetto alla primitiva posizione logicista, dell'intuizionismo brouweriano: né la filosofia né qualunque altra scienza, e quindi in particolare non la logica, possono fondare la matematica, che viceversa si costituisce a partire da un'intuizione generale e fondamentale di tipo immediatamente apprensibile da parte del soggetto. È semmai sulla *controparte linguistica* della matematica data come fatto intuitivo originario che può costituirsi una logica.

Detto questo, veniamo all'altro punto per noi interessante. Per Russell, o comunque per i logicisti in generale, è il problema della *forma* che viene in primo piano. Un unico grande sistema formale, la Grande Logica, deve rendere conto di tutto il contenuto della matematica; qui la logica viene soprattutto intesa, in senso globale, come espressione della massima *generalità*, o se si preferisce, come espressione della massima possibilità di generalizzazione. L'esistenza degli oggetti matematici diventa quindi automaticamente qualcosa di garantito dalla *possibilità di definizione* nell'ambito di

un tale sistema la cui correttezza e completezza formale (intese in senso intuitivo) conferiscono e determinano il contenuto stesso della matematica. In opposizione a questo «contenutismo formale» sono accomunati – pur se con sfumature e concezioni di fondo peraltro notevolmente diverse – tanto i formalisti hilbertiani e Zermelo quanto gli intuizionisti. Secondo queste concezioni si tratta di esplicitare dei *contenuti matematici* che vengono tuttavia visti in modi completamente diversi nei due casi, e precisamente come corrispettivi di insiemi di enunciati, siano essi assiomi o regole (per la deduzione di altri enunciati da enunciati dati) nel caso di Hilbert e (in senso lato) di Zermelo, o come corrispettivi di atti mentali, di costruzioni nel caso di Brouwer (e, con le opportune precisazioni, anche di Borel e Poincaré). Nel primo caso ne viene che l'esistenza si risolve in *coerenza* di insiemi di enunciati o sistemi assiomatici, mentre nel secondo caso la risoluzione dell'esistenza degli oggetti matematici sta proprio, ovviamente, nella *possibilità di costruzione*. E in fin dei conti è proprio sul contenuto e il tipo di quella «possibilità» che si differenziano tra loro le varie posizioni costruttiviste.

Qual era a questo riguardo la posizione dei protagonisti indiscussi dell'indagine sui fondamenti della geometria sul finire del secolo, i rappresentanti di quella scuola italiana che Russell stesso riconosceva come i propri maestri? Salvo che Peano, Cesare Burali-Forti (1861-1931) e Mario Pieri essi sono stati nominati solo di sfuggita. E in effetti, se un altro nome si deve aggiungere è soltanto quello di Alessandro Padoa che fra il 1900 e il 1904 presenta una serie di risultati sulla definibilità dei concetti di cui parleremo nel paragrafo 2 del capitolo VI. Un altro nome *vorremmo* poter aggiungere, quello di Giovanni Vailati la cui prematura scomparsa ha privato indubbiamente l'ambiente culturale italiano del periodo di un possibile apporto certamente significativo; e diciamo «vorremmo» per due opposti motivi: da una parte perché egli era una fra le poche figure con una visione *globale* della cultura, con interessi assai vari e certamente non «provinciale»; dall'altra perché Vailati in effetti pur facendo parte della scuola di Peano non si interessò che marginalmente di logica. Va aggiunto che nel 1905 vedeva la luce quella *Logica come scienza del concetto puro* di Benedetto Croce (1866-1952) che parrebbe offrire secondo alcuni – se non l'unica – almeno la ragione principale del totale decadimento di

una scuola pur apparentemente tanto «vitale» come quella peaniana.

Non è qui il caso di valutare e analizzare le influenze e i condizionamenti che l'opera complessiva di Croce ha avuto sulla cultura, anche scientifica, italiana. Non possiamo tuttavia esimerci dal notare come ben difficilmente si riesca a trovare in un'altra opera un repertorio così vasto e nutrito di pretenziose inesattezze, superficialità, di vere e proprie insulsaggini per quanto riguarda la logica «formalistica» – come Croce la chiama – e i vari tentativi di riformarla «... il più solenne dei quali è la già ricordata Logica matematica, denominata anche calcolatoria, algebrica, algoritmica, simbolica, nuova analitica, calcolo logico o Logistica».[25] E mentre in tutte le *culture* europee (basti pensare a quella francese, o inglese, o tedesca) tutta una serie di studiosi affrontava con serietà e rigore i profondi problemi che erano stati sollevati dalle antinomie, dalle diverse concezioni che si dividevano il campo intorno al problema dell'esistenza degli enti matematici, da noi invece il saggio di Pescasseroli, il genio universale della filosofia italiana, decideva che «se come *scienza del pensiero* la Logistica è cosa risibile, degna veramente dei cervelli che l'hanno costruita [e abbiamo avuto occasione di nominare in queste pagine alcuni di quegli inetti e limitati cervelli che tale brutta azione avevano commesso]... non è poi nostro assunto esaminarla in quanto formulario provvisto di *pratica utilità*; e su questo punto ci restringiamo a insistere sopra una sola e assai semplice osservazione».

Da Leibniz in poi, osserva sostanzialmente Croce, «questi nuovi congegni sono stati offerti sul mercato: e tutti, sempre li hanno stimati troppo costosi e complicati, cosicché non sono finora en-

[25] Ovvio che «formalistica» vuol essere un dispregiativo il cui impiego Croce ritiene di poter giustificare con dovizia di motivazioni. A parte il fatto che per lui «ora è il bel tempo dei Peano, dei Boole [sic!], dei Couturat» e che per lui «la stessa dottrina della *quantificazione del predicato*» è stata un po' il «lievito» della «riforma» della logica, è sconsolante vedere come uno scrittore così acuto, brillante e profondo possa cadere semplicemente nel ridicolo per una totale mancanza di informazione spicciola, banale (che ovviamente diventa grossa responsabilità culturale quando, ed è questo il caso, ha una chiara origine dogmatica). In particolare tutto il terzo capitolo della sezione seconda della prima parte di quest'opera di Croce è un coacervo di tali amenità! È chiaro quindi che se ci occupiamo di Croce in questo contesto è soltanto per l'indubbia influenza e per il grosso condizionamento che egli ha esercitato (e parzialmente continua a esercitare) sulla cultura italiana.

trati né punto né poco nell'uso. *Vi entreranno nell'avvenire? La cosa non sembra probabile*, e, ad ogni modo, è fuori dalla competenza della filosofia e appartiene a quella della pratica riuscita: da raccomandarsi, se mai, a commessi viaggiatori che persuadano dell'utilità della nuova merce e le acquistino clienti e mercati. Se molti o alcuni adotteranno i nuovi congegni logici, questi avranno provato la loro grande o piccola utilità. Ma la loro nullità filosofica rimane, fin da ora, pienamente provata» [corsivo nostro].

Scusi il lettore la lunga citazione che tuttavia ci sembrava troppo significativa e paradigmatica. E naturalmente solo su questo si potrebbe parlare a lungo, ma tempo e spazio possono essere impegnati, a nostro parere, in modo assai più proficuo sicché ci si può limitare a osservare – a parte l'impressionante, quasi profetico potere di previsione di cui Croce fa sfoggio in questo passo – che probabilmente l'accenno «economicista» è stato stilato di getto, avendo presente il «mercato» di Pescasseroli, o, a essere compiacenti, quello italiano. Per quanto riguarda infine il rapporto della logica con la filosofia (e sempre, per essere brevi, su pure basi pragmatiche) va invece tutto bene: purché Croce a quella «filosofia» premetta un «mia»: perché allora è certo che la logica matematica non ha nulla a che fare con la *sua* filosofia.

Purtroppo, come già ricordato, la situazione sta in termini ben più gravi e complessi, perché non può imputarsi solo a Croce la decadenza della scuola logica italiana; Croce piuttosto intervenne nel decretare per le scienze in generale – e quindi per la logica in particolare – l'esclusione dall'«area culturale», relegandole a mere manipolatrici di pseudoconcetti. Ma tutto ciò avvenne con la decisiva e colpevole complicità degli scienziati italiani.

Per quanto in particolare riguarda la logica, il suo declino inizierà paradossalmente proprio con Peano, con la sua «sordità», immediatamente ereditata dai suoi «scolari», verso una collocazione più generale dell'enorme problematica che lui stesso aveva contribuito a sollevare nel mondo. E ciò sembra dovuto alla sua tendenza a «chiudere» un sistema, una ricerca, un'impresa, se necessario forzatamente, invece di spingersi fino alle estreme conseguenze *sperimentali* di nuovi tentativi, principi, metodi: è forse un malinteso senso di «onestà intellettuale» che lo porta a escludere dai suoi discorsi ogni considerazione che non sia, come lui stesso dice, di «stretta pertinenza matematica». Peano, che già dovette

combattere in patria contro coloro che ritenevano frutto di «senilità» la sua produzione e i suoi interessi logici, ci appare inevitabilmente conservatore e provinciale non appena ci si affacci oltre i confini italiani (anche se – paradossalmente – è proprio all'estero che la sua influenza positiva si fece sentire); e ciò non solo perché come ebbe a scrivere nel 1915 «la logica matematica, utile nei ragionamenti matematici (ed in questo senso io ne feci uso) interessa pure la filosofia», portando poi Louis Couturat come esempio in proposito; ma anche soprattutto perché sempre nello stesso articolo (*Importanza dei simboli in matematica*) afferma testualmente: «...se giuste sono dunque le critiche di Eugenio Rignano contro coloro che *considerano la logica matematica quale scienza in sé*, i cui lavori, è verissimo, sono spesso poco proficui; invece più non lo sarebbero all'indirizzo di coloro... che considerano la logica matematica come uno *strumento utile* per risolvere questioni *matematiche resistenti ai metodi comuni*» [corsivi nostri]. Come unico commento basterà qui ricordare al lettore che solo due anni prima era comparsa la prima edizione dell'ultimo volume dei *Principia.*

Del resto le stesse formule di «disimpegno» di Peano sono estremamente illuminanti in questo senso. Quando, ad esempio, a proposito del paradosso di Richard, egli dichiara emblematicamente che «exemplo de Richard non pertinet ad mathematica sed ad linguistica» ciò rappresenta per lui un modo di liquidare una questione che non solo non lo *interessa* come *matematico* perché ha varcato i limiti della sua competenza (ché qui ancora potremmo parlare di «onestà intellettuale») ma che respinge come uomo di cultura perché è al di là della sua stessa apertura mentale. Si pensi soltanto che la sua affermazione sopra riportata agirà non come freno, bensì come potente catalizzatore, come lievito, per un autore ben diversamente disposto, quel F.P. Ramsey di cui avremo occasione di parlare in un prossimo paragrafo.

Senza neppure ricordare le parole di Zermelo sopra viste non può non colpire la singolare vicenda intellettuale di un uomo come Peano – indubbiamente grande matematico e grande logico – destinato a fornire spunti di apertura a Russell, a Ramsey, a Couturat, a tutto un mondo culturale insomma, e tuttavia decisamente *bloccato* a livello di una propria elaborazione culturale autonoma, di un inserimento dei suoi stessi metodi e problemi in un più ampio contesto. Fatalmente, questa visione ristretta venne eredi-

tata – grazie anche a una robusta dose di puro «accademismo» che alla nostra cultura universitaria non ha certo mai fatto difetto – dai suoi «scolari». A leggere articoli, conferenze, lavori di ogni tipo, degli allievi di Hilbert (che non è certo matematico di levatura inferiore a quella di Peano), ad esempio di un Paul Bernays, di un Wilhelm Ackermann, di un John von Neumann (1903-57) non si troveranno *mai* locuzioni come «il grande Hilbert» o «come ha definitivamente mostrato il nostro maestro Hilbert» o altre di questo tenore che con tutta tranquillità e regolarità *abbondano* invece negli scritti di un Burali-Forti, o di un Padoa, o di un Vacca (ovviamente con la sostituzione Hilbert-Peano): si tratta, nel primo caso di discutere, approfondire, rielaborare, superare un risultato che o si sa essere di Hilbert o si riconosce come tale per dovuta correttezza; nel secondo invece si tratta di sottolineare l'importanza della scuola e del «maestro», quindi di aggiungere modestamente, con le dovute cautele, qualche briciola. È chiaro che la causa di tale situazione non può farsi risalire a Peano, ma a tutto un ambiente che del resto solo oggi (1992) *tende* in generale a modificarsi.

Riteniamo che le seguenti parole di Burali-Forti, tratte dalla seconda edizione (1919) della sua *Logica matematica*, mostrino con sufficiente chiarezza come la cultura italiana – anche specifica – dell'epoca si fosse estraniata dai grandi temi dibattuti nel resto del mondo dalle varie scuole e si fosse già rinchiusa in un provincialismo masochistico senza respiro né sbocchi. Burali-Forti dunque si augura che la sua fatica possa avere migliore fortuna di altre e nel contempo dichiara che volendo «presentare al lettore la logica simbolica *in azione*, cioè come *strumento* di *analisi* e di *scrittura abbreviata* [e qui è difficile non dar ragione a Croce!] mi son potuto risparmiare l'enorme [!] fatica di stabilire un sistema completo di proposizioni primitive; *tanto più che è dubbio se tale enorme fatica potrebbe essere coronata da successo, visto che i tentativi fatti finora sono ben lontani dall'aver dato risultati esaurienti*» [corsivo nostro]. Dove appunto è chiaro che gli italiani non ricercavano o sperimentavano ma o avevano «risultati esaurienti» oppure niente! E la nota finale mi sembra addirittura proponga quel triste e speriamo ormai lontano e superato anelito autarchico-nazionalista che ben altro peso avrebbe dovuto purtroppo avere sulla nostra vita politica e sociale: «E a proposito di *merce estera* – afferma Burali-Forti – con-

viene far notare come, specialmente dagli italiani (!), si citi e si usi il caotico e impreciso sistema geometrico dell'Hilbert, quasi non esistessero i sistemi semplici, chiari e precisi (ma sono italiani!) e ben superiori a quello dell'Hilbert, di M. Pieri».

Si potrebbe continuare a lungo in questo discorso: ma non ci sembra questa la sede più opportuna per una compiuta analisi storica di un fatto tutto sommato così particolare. Quello che ci preme sottolineare è che per tutto un complesso di ragioni, dopo il primo decennio del secolo la scuola italiana non costituisce più un interlocutore nel dibattito generale sui fondamenti della matematica e sul rinnovamento della logica.

3. NUOVI ORIZZONTI: GLI ANNI VENTI

Lasciato da parte l'ambiente italiano, ormai decisamente emarginato, si può osservare che il panorama della ricerca negli anni venti si presenta estremamente vario: i problemi sul tappeto sono molti e complessi, dopo questa prima, rinnovata fase di indagine sulla logica e sui fondamenti della matematica. Punto di riferimento obbligato, ancora agli inizi degli anni venti, sono sempre i *Principia mathematica* (che vedranno la seconda edizione fra il 1925 e il 1927) riguardati – oltre che, ovviamente, come manifesto della scuola logicista – come momento operativo di analisi logica di tutta la matematica, con la preoccupazione centrale di evitare i paradossi. Abbiamo visto quale fosse stato il prezzo pagato da Russell e Whitehead per questo ambizioso programma: a parte locali e contingenti lacune di chiarezza o rigore, si trattava più intrinsecamente della complessità e pesantezza della teoria dei tipi ramificati e delle necessarie e assai impegnative assunzioni che gli autori dei *Principia* furono costretti a fare per garantire la ricostruzione della matematica: assioma di riducibilità, assioma moltiplicativo, assioma dell'infinito. La natura esplicitamente esistenziale di questi assiomi faceva emergere in modo naturale domande di fondo circa la natura stessa della logica nel cui ambito doveva, secondo la tesi logicista dei *Principia*, essere riportata tutta la matematica.

Detto in modo diverso, quello che in certo senso possiamo assumere un po' come *leitmotiv* della ricerca logica e fondazionale di questo nuovo periodo è il problema che può esprimersi con le do-

mande: cos'è la logica? qual è il suo linguaggio? e quali rapporti ha la logica con i contenuti matematici? Questi interrogativi sorgevano immediati sia all'interno dei diversi programmi fondazionali (non solo quello logicista, ma anche quello più genericamente insiemistico, per il quale vitale era il problema di determinare che cosa significasse *proprietà definita*) sia in contesto più specificamente logico, una volta che si cercasse di analizzare frammenti del linguaggio dei *Principia* indagandone le capacità e i limiti espressivi. Tanto i lavori dell'americano Emil Leon Post relativi alla logica enunciativa intesa come sistema formale autonomo, quanto la proposta di logiche «non classiche» avanzata da Clarence Irving Lewis e da alcuni esponenti della scuola polacca, come infine la canonizzazione dei vari calcoli data nel 1928 da Hilbert e Ackermann possono infatti essere riguardati come momenti diversi di questa seconda esigenza.

Più articolata la situazione per quanto riguarda la ricerca fondazionale e i suoi rappoti con l'indagine logica. Per la scuola intuizionista, oltre ai numerosissimi interventi di Brouwer di chiarimento e riedificazione di una matematica intuizionista o di polemica contro la scuola formalista, questo periodo vede nascere un primo e interessantissimo tentativo di formalizzazione della «logica» (e della matematica) intuizioniste; questo tentativo oltre ad anticipare risultati che matureranno compiutamente all'inizio degli anni trenta, renderà possibile l'instaurarsi di un primo concreto e fruttuoso confronto fra le idee intuizioniste e quelle delle altre scuole.

Si assiste invece ad una sorta di «ridimensionamento» della scuola logicista che, pur difesa ancora con vigore da Rudolf Carnap proprio verso il 1930, giunge sostanzialmente a un punto morto tanto a causa di critiche «esterne» quanto per i profondi ripensamenti che autori qualificantisi essi stessi come logicisti, primo fra tutti F.P. Ramsey, vengono formulando in questi anni.

Questo decennio è invece particolarmente fecondo per la scuola formalista che vede Hilbert e tutta una schiera di collaboratori, fra i quali Ackermann, von Neumann, Bernays e altri (in particolare J. Herbrand in Francia) impegnati nell'esposizione ed elaborazione della piattaforma teorica definitiva del programma formalista e nella presentazione di numerosi, interessanti e assai pro-

mettenti risultati parziali che portano lo stesso mondo matematico alla sostanziale pur se implicita convinzione della completa e pressoché certa attuazione del programma hilbertiano.

È nel contesto della assiomatizzazione della teoria degli insiemi che però vanno ricercate le motivazioni ultime riguardanti i diversi tentativi di isolare una concezione della logica più adeguata al suo ruolo fondazionale. Si ricorderà come Zermelo, formulando il suo assioma caratteristico di isolamento, avesse fatto ricorso al concetto di *proprietà definita* cercando mediante esso di doppiare gli scogli da una parte di una embrionale tipizzazione, dall'altra di paradossi semantici come quello di Richard e di König. Era la prima volta che in modo esplicito si presentava la necessità di determinare un concetto chiaramente logico per scopi dichiaratamente ed esclusivamente matematici. L'indeterminatezza che, come vedremo, fu subito sottolineata, del concetto introdotto da Zermelo, non tardò a portare in primo piano il problema generale di che cosa fosse un linguaggio formale, di come valutare le sue capacità espressive ed infine di come presentarlo senza presupporre fatti insiemistici controversi. Come vedremo, i dibattiti coinvolsero Zermelo stesso, Fraenkel, von Neumann e soprattutto Skolem, sul problema della categoricità della teoria degli insiemi formalizzata e portarono ad analisi ravvicinate delle distinzioni fra linguaggi del primo e del secondo ordine, fra sintassi e semantica, fra metodi finitisti e non, che contraddistinguono la discussione negli anni venti.

3.1 *La rottura della Grande Logica. Logiche non classiche*

Abbiamo visto come alla sistemazione dei *Principia* fossero state sollevate notevoli critiche. Quelle cui finora abbiamo accennato, e che prenderemo in esame più dettagliato in paragrafi successivi, riguardavano tuttavia quasi esclusivamente la concreta realizzazione del sistema logico dei *Principia* in rapporto al programma logicista professato dai suoi autori (in particolare da Russell). Oltre però a rappresentare un potente stimolo alla discussione fra scuole e alla ricerca in questa direzione, i *Principia* diedero lo spunto a tutta un'altra serie di riflessioni, solo apparentemente più locali, e che sfociarono in varie proposte alternative che avranno un enor-

me peso nello sviluppo successivo della logica in sé e per sé. Le indagini in questa seconda direzione possono farsi cominciare con il *Survey of symbolic logic* (*Panorama di logica simbolica*, 1918) dell'americano Lewis e proseguono in effetti per tutti gli anni venti, con notevoli contributi di autori americani e soprattutto degli studiosi della cosiddetta scuola polacca. In quest'ordine di idee può situarsi anche la «canonizzazione» della frattura della Grande Logica di Frege e Russell che si ebbe nel 1928 con i *Grundzüge* di Hilbert e Ackermann, cui sopra abbiamo accennato.

Occupiamoci, per cominciare, di quell'aspetto della questione che può riassumersi nell'osservazione che già nel 1920 faceva Post quando scriveva che «nella teoria generale della logica costruita da Whitehead e Russell... per fornire la base per tutta la matematica, esiste una certa sottoteoria che è unica nella sua semplicità e precisione e, malgrado tutte le altre parti dell'opera si fondino su questa sottoteoria, ... è completamente indipendente da quelle». Post alludeva a quella che oggi viene chiamata *logica delle proposizioni* ossia a quella parte del sistema logico dei *Principia* che riguarda semplicemente le connessioni inferenziali che possono stabilirsi considerando come «dati iniziali» proposizioni non analizzate (e non funzioni proposizionali) quando esse vengano variamente collegate fra loro tramite connettivi logici.

L'interesse che muoveva Post in questa sua osservazione, le motivazioni che lo spingevano a proporre le generalizzazioni cui accenneremo erano per sua dichiarazione esplicita *puramente formali*, sicché le estensioni cui egli giungeva sono semplici e naturali estensioni suggerite dalla struttura stessa della teoria così isolata, senza alcuna preoccupazione per le eventuali interpretazioni logiche che tali risultati potevano o meno avere. Quest'atteggiamento consentiva a Post, tra l'altro, di cogliere ed esplicitare la distinzione fra il linguaggio della teoria che è oggetto di studio e il linguaggio nel quale si conduce lo studio della teoria stessa tracciando così una consapevole linea di demarcazione fra teoria e *metateoria* – per usare una terminologia oggi corrente – o se si preferisce, nel caso particolare, fra *logica* e *metalogica*: «Vogliamo qui porre l'accento sul fatto che i teoremi di questo lavoro vertono *sulla* logica ma *non sono inclusi* in essa». Era una ripresa, nell'ambito della tradizione logicista, dell'approccio algebrico dei postulazionisti, cui Post si rifaceva e che Lewis aveva chiaramente posto a confronto

col metodo logicista nel *Survey*. Questo atteggiamento «metamatematico» sarà caratteristico della posizione hilbertiana mentre il discorso metamatematico stesso – come abbiamo già anticipato – diverrà oggetto di studio approfondito per la scuola polacca.

Naturalmente la parte proposizionale della logica era nota almeno dal tempo di Frege come sistema autonomo, e in effetti già da Boole come interpretazione del proprio sistema formale; e anzi tipiche ricerche di possibilità di riduzione delle assunzioni e dei termini fondamentali avevano portato Henry Maurice Sheffer (1883-1964) nel 1913 a ridurre a uno solo i connettivi proposizionali e Jean Nicod (1893-1924) nel 1917 a ridurre a un solo assioma la base assiomatica proposizionale. Ma tutto ciò era sempre inteso come semplificazione operativa o di principio *all'interno* del sistema di Russell (o, nel caso di Frege, come primo passo verso la Grande Logica). Merito di Post è invece quello di aver isolato il sistema di logica proposizionale (o enunciativa, come anche diremo) come una teoria assiomatica e deduttiva autonoma, assumendo come unico oggetto di indagine, senza altro contesto, una teoria costruita su una porzione estremamente limitata del linguaggio «universale» dei *Principia*, e distinguendo i due livelli linguistici della teoria e della metateoria come piani su cui muoversi *separatamente*. Il linguaggio proposizionale, il «minimo» linguaggio logico disponibile, dal punto di vista classico, ha un'importanza centrale che sarebbe chiaramente emersa una volta che, affrontando problemi di decisione, si vide che i quantificatori introducevano ineliminabili riferimenti di tipo infinitario che non si davano nel linguaggio proposizionale. La logica proposizionale costituiva così un nucleo importante di chiaro significato finitista, una sorta di paradigma cui ricondurre eventualmente ogni altro linguaggio più forte. Questo malgrado fosse del tutto chiara la povertà espressiva dei linguaggi proposizionali, come sottolineava lo stesso Post quando auspicava un prosieguo del suo lavoro su teorie logiche espresse con linguaggi più «sottili e ricchi di informazione».

Post dà una versione estremamente precisa della logica proposizionale dei *Principia* tanto dal punto di vista assiomatico, quanto dal punto di vista semantico. Sul piano sintattico Post assumeva come nei *Principia*, i due soli connettivi di negazione ($\neg$) e disgiunzione ($\vee$), sulla base dei quali, a partire dalle lettere proposizionali, si costituiscono tutte le rimanenti formule del linguaggio;

sul piano semantico dava una versione rigorosa delle tavole di verità a due valori e ne mostrava i molteplici usi.

Il sistema proposizionale di Post è caratterizzato da un certo numero finito di assiomi che ovviamente sono tautologie, e da due regole di inferenza: il *modus ponens* o regola di separazione, che dai teoremi $\mathscr{A}$ e $\mathscr{A} \rightarrow \mathscr{B}$ permette di derivare il teorema $\mathscr{B}$

$$\frac{\vdash \mathscr{A}, \vdash \mathscr{A} \rightarrow \mathscr{B}}{\vdash \mathscr{B}}$$

e la regola di sostituzione, secondo la quale a ogni lettera indicante una proposizione (lettera o variabile proposizionale) si può sostituire qualunque altra formula del linguaggio (ottenendo anche in questo caso teoremi da teoremi). Il calcolo è organizzato in modo che a partire dagli assiomi, ossia da tautologie, si possono ottenere, applicando le regole di deduzione, teoremi che si rivelano a loro volta tautologie. Ciò si esprime dicendo che il sistema è *valido* (ossia dà luogo solo a teoremi che sono *logicamente validi*). Si poneva in modo naturale anche la domanda inversa che già nel 1918 P. Bernays aveva posto e risolto nella sua *Habilitationschrift* (resa nota solo più tardi): il calcolo è in grado di farci ottenere come teoremi *tutte* le tautologie? *Completezza semantica* (o, in terminologia più recente, *adeguatezza*) è il nome di questa proprietà del calcolo, che, nel caso ne goda, viene appunto detto (semanticamente) completo o adeguato.[26] Post dimostra che il sistema proposizionale è completo in questo senso ma allarga l'indagine ad altri tipi di completezza, introducendo due altre nozioni. Secondo la prima (detta oggi *completezza nel senso di Post*) un sistema è completo se ogni sua formula diventa dimostrabile ogniqualvolta noi aggiungiamo come assioma una qualunque formula non dimostrabile nel sistema; la seconda è quella che noi oggi diciamo *completezza funzionale* (e a questa proprietà Post riserva il nome di completezza) e consiste nel fatto che ogni funzione di verità può essere scritta in termini dei due connettivi primitivi da lui assunti.

[26] Ove non sorgano confusioni diremo anche semplicemente completezza invece di completezza semantica; la ragione della necessità, almeno in certi contesti, della qualificazione, consiste nel fatto che come si ricorderà abbiamo accennato nel capitolo III a un altro tipo di completezza che viene detta *sintattica*.

Nella sua dimostrazione, che sfrutta particolari forme normali, Post prova anche che le tavole forniscono una *procedura di decisione* per la logica proposizionale e cioè un metodo *meccanico* grazie al quale, data una qualunque formula proposizionale, è possibile stabilire in un numero finito di passi se tale formula è un teorema oppure no. Post dimostra inoltre che il sistema da lui considerato è coerente, ossia che in esso non è possibile dimostrare contemporaneamente una formula $\mathscr{A}$ e la sua negazione $\neg \mathscr{A}$ ma introduce anche un'altra nozione di coerenza, detta talora *coerenza nel senso di Post*: un calcolo che contiene variabili proposizionali è coerente in questo senso se in esso non può essere dimostrata una singola variabile proposizionale. Post dimostra che il sistema da lui considerato è coerente anche in questo senso. Si noti che i due concetti coincidono per sistemi che ammettano la negazione come connettivo.

L'importanza di questi risultati stava soprattutto nelle prospettive che aprivano. Per Post non è dato a priori né lo stock dei connettivi fondamentali, né la scelta delle regole e degli assiomi del calcolo né – come vedremo – il numero stesso dei possibili valori di verità con le associate tavole per i connettivi. Esistono diverse alternative, che vanno indagate sistematicamente, ed è per questo che Post pone e risolve problemi come quello della completezza funzionale e semantica, definisce diverse nozioni di completezza e coerenza, ecc. In altre parole – anche a prescindere dalle generalizzazioni diverse da quella classica, di cui parleremo più avanti – l'oggetto di Post non è una specifica formulazione della logica proposizionale classica, ma l'intera varietà delle diverse formulazioni e dei loro mutui rapporti. È l'atteggiamento – chiaramente antagonista a quello di Russell e Witehead – degli algebristi della logica che si afferma, prospettando un contesto più generale entro il quale porre la logica dei *Principia*. Non è azzardato dire che queste ricerche di Post – assieme a quelle più o meno coeve di Łukasiewicz e della sua scuola – segnano l'inizio di uno studio autonomo dei sistemi proposizionali e non stupisce che Post – ancora come Łukasiewicz – prospetti la possibilità di logiche diverse da quella classica.

Una prima generalizzazione riguarda da una parte sistemi con un numero finito qualunque di connettivi, dall'altra sistemi con un numero finito qualunque di assiomi e/o regole di inferenza.

Una seconda generalizzazione riguarda invece la considerazione di «tavole di verità» (che vengono ora dette più appropriatamente «matrici») con un numero finito di valori, cioè logiche proposizionali con un numero finito di «valori di verità» maggiore o uguale a 2. Da un punto di vista intuitivo ciò comportava evidentemente prendere in esame, accanto ai due tradizionali valori di verità Vero e Falso, anche altri valori di verità come potrebbero essere ad esempio «incerto», «indifferente», «indeterminato» e simili: in linguaggio moderno, *logiche polivalenti*, invece della classica logica bivalente. Va ribadito tuttavia che per Post queste generalizzazioni si presentano come naturali illustrazioni del concetto generale di logica proposizionale e non comportano minimamente la preoccupazione di una eventuale giustificazione logica che le ponga in alternativa con la logica classica come più o meno nello stesso giro d'anni – con motivazioni filosofiche diverse – facevano Nikolai A. Vasil'ev (1880-1940), con la sua logica immaginaria modellata sulla geometria di Lobačevskij, e Łukasiewicz (sul quale torneremo fra breve).

Come vedremo più avanti, dopo la logica proposizionale altre logiche più espressive sarebbero state isolate all'interno dei *Principia*, prima fra tutte quella elementare. Una sistemazione paradigmatica di questa frammentazione si ritrova nei *Grundzüge* di Hilbert e Ackermann. Il criterio di suddivisione era già stato implicitamente accennato nei *Principia* e si basa sulle possibilità espressive dei linguaggi anche in relazione al tipo di analisi che consentono di fare del contenuto che sono destinati a esprimere. La prima grande distinzione avviene quindi tra *linguaggio enunciativo*, i cui elementi base sono proposizioni non analizzate e connettivi, e *linguaggio predicativo* in cui le proposizioni sono pensate analizzate nelle loro componenti *predicativa* (o più in generale *relazionale*) e *soggettiva*; oltre ai connettivi proposizionali il linguaggio predicativo comprende gli operatori logici di quantificazione. Nell'ambito del linguaggio predicativo interviene una distinzione ulteriore a seconda che si convenga di applicare la quantificazione alle sole variabili individuali (potendo così esprimere nel linguaggio frasi come, ad esempio «per tutti gli individui...» o «esiste almeno un individuo tale che...») oppure si intenda estendere la quantificazione anche a variabili predicative (sicché diventano esprimibili frasi quali «per tutte le proprietà (di individui)...» o «esiste alme-

no una proprietà (di individui) tale che...»). Si ottiene così il *linguaggio predicativo del primo ordine* (quantificazione limitata) o del *secondo ordine* (quantificazione estesa). Dovrebbe essere chiaro com'è possibile continuare in questa gerarchia; va detto però che raramente si avverte la necessità di linguaggi di ordine superiore al secondo e che in generale il linguaggio del primo ordine, pur essendo una porzione molto limitata del linguaggio comune, è sostanzialmente sufficiente per esprimere almeno la maggior parte delle teorie matematiche interessanti e i grandi dibattiti che si ebbero negli anni venti riguardano piuttosto la contrapposizione dei linguaggi del primo e del secondo ordine.

Nell'impostazione di Hilbert e Ackermann, un linguaggio proposizionale consta di un insieme per lo più infinito numerabile di *lettere proposizionali* (destinate a essere interpretate su proposizioni, o meglio su valori di verità) p, q, r, s ..., dei connettivi $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, di simboli ausiliari quali le parentesi, la virgola ecc. Per costruire su questo «alfabeto» il sistema deduttivo del calcolo proposizionale, occorre definire le formule, gli assiomi e dare le regole di inferenza. Intuitivamente le formule sono, fra tutte le espressioni che possono costruirsi componendo linearmente segni dell'alfabeto, quelle che vogliamo considerare come sintatticamente ammissibili (è chiaro peraltro che esse saranno scelte in modo tale da risultare «significanti» anche da un punto di vista semantico, sotto opportune interpretazioni standard); gli assiomi sono, fra le formule, quelle che noi scegliamo come elementi iniziali del procedimento dimostrativo.

L'insieme delle formule è definito in modo induttivo, secondo una procedura la cui rilevanza Hilbert aveva sottolineato nel suo articolo del 1904:

1) ogni lettera proposizionale è una formula;
2) se $\mathscr{A}$ e $\mathscr{B}$ sono formule allora anche $(\neg\ \mathscr{A})$, $(\mathscr{A} \wedge \mathscr{B})$, $(\mathscr{A} \vee \mathscr{B})$, $(\mathscr{A} \rightarrow \mathscr{B})$, $(\mathscr{A} \leftrightarrow \mathscr{B})$ sono formule;
3) nient'altro è una formula.[27]

[27] La 2) mostra come l'introduzione di ogni connettivo comporti l'uso di parentesi (ad esempio da $\mathscr{A}$ e $\mathscr{B}$ $(\mathscr{A} \vee \mathscr{B})$). Questo è reso necessario, in assenza di clausole sulla forza dei connettivi, dall'esigenza di avere – come nelle espressioni algebriche che il lettore avrà incontrato nelle scuole inferiori – una lettura unica, un'univoca decomposizione in sottoformule. Detto questo, vista la grande com-

Sono così disponibili tanto un metodo di *dimostrazione* per induzione per provare fatti sulla totalità delle formule quanto un metodo di definizione per ricorsione, per definire relazioni e operazioni su formule. Discorso analogo si può fare per il calcolo logico ed il concetto di teorema.

Gli assiomi sono 15 divisi in cinque gruppi, ognuno dei quali regola il comportamento di un connettivo.

Assiomi dell'implicazione

1) $p \rightarrow (q \rightarrow p)$
2) $(p \rightarrow (p \rightarrow q)) \rightarrow (p \rightarrow q)$
3) $(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow (p \rightarrow r))$

Assiomi della congiunzione

4) $p \wedge q \rightarrow p$
5) $p \wedge q \rightarrow q$
6) $(p \rightarrow q) \rightarrow ((p \rightarrow r) \rightarrow (p \rightarrow q \wedge r))$

Assiomi della disgiunzione

7) $p \rightarrow p \vee q$
8) $q \rightarrow p \vee q$
9) $(p \rightarrow r) \rightarrow ((q \rightarrow r) \rightarrow (p \vee q \rightarrow r))$

Assiomi dell'equivalenza

10) $(p \leftrightarrow q) \rightarrow (p \rightarrow q)$
11) $(p \leftrightarrow q) \rightarrow (q \rightarrow p)$
12) $(p \leftrightarrow q) \rightarrow ((q \rightarrow p) \rightarrow (p \rightarrow q))$

Assiomi della negazione

13) $(p \rightarrow q) \rightarrow (\neg q \rightarrow \neg p)$
14) $p \rightarrow \neg \neg p$
15) $\neg \neg p \rightarrow p$

plicazione che il seguire queste norme sulle parentesi comporterebbe, conveniamo di scrivere le formule di cui avremo bisogno (ad esempio gli assiomi qui di seguito) eliminando le parentesi ogni qual volta questo non comporti equivoci nella lettura. Si tratta di una convenzione *metalinguistica* che non altera la definizione di formula.

Come regole assumiamo la regola di separazione e la regola di sostituzione di cui si è detto sopra.[28]

Teorema logico sarà ogni formula ottenibile dagli assiomi applicando un numero finito di volte le regole suddette. È chiaro allora che l'insieme dei teoremi risulta definito induttivamente come la più piccola famiglia di formule contenente gli assiomi e chiusa rispetto alle regole. Ciò ci permette di utilizzare metodi di definizione e di dimostrazione per induzione e – tenuto conto del carattere puramente sintattico, concreto, della definizione di assioma e di regola – ci mostra come la nozione di teorema logico risulti in via di principio definita in modo concettualmente più semplice di quella di legge logica intesa in senso semantico. In modo analogo questa contrapposizione tra aspetto sintattico e semantico si presenterà nel caso della logica del primo e del secondo ordine, acquistando un carattere molto più significativo in vista della natura decisamente infinitaria delle rispettive semantiche.

Per la logica del primo ordine avremo come possibile naturale estensione la seguente. Il linguaggio comprende: *a*) un insieme infinito numerabile di variabili individuali x_1, x_2, x_3, ... x_n, ...; *b*) le lettere proposizionali, che verranno ora indicate per comodità con P^0_0, P^0_1, P^0_2, ...; delle lettere predicative (almeno una) indicate con P^k_i dove i (≥ 0, finito) è un indice di differenziazione e k (≥ 1, finito) indica il numero di posti della lettera predicativa (si potranno considerare le lettere proposizionali come lettere predicative «de-

[28] Come sistemazione assiomatica abbiamo in realtà riportato quella di Hilbert e Bernays, *Grundlagen der Mathematik* (*Fondamenti della matematica*, 2 volumi; 1934-1939). Naturalmente questo sistema è equivalente a quello originale di Hilbert-Ackermann e ad altri allora correnti e che si differenziavano fra loro e per le diverse assunzioni «linguistiche» (numero e scelta dei connettivi) e per le diverse assunzioni «deduttive» (assiomi e regole). Attualmente si preferisce in generale dare tali sistemi a livello *metalinguistico*, usando cioè nella enunciazione degli assiomi variabili metalinguistiche che possono essere immaginate sostituite in modo uniforme da formule linguistiche qualunque. È una semplificazione introdotta da von Neumann che sostituisce agli assiomi *schemi* d'assiomi. Questo permette di eliminare la regola di sostituzione. Un esempio di tali sistemi, basato su due connettivi e tre schemi di assiomi, è il seguente:

$$\mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{A})$$
$$(\mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{C})) \rightarrow ((\mathcal{A} \rightarrow \mathcal{B}) \rightarrow (\mathcal{A} \rightarrow \mathcal{C}))$$
$$(\neg \mathcal{B} \rightarrow \neg \mathcal{A}) \rightarrow ((\neg \mathcal{B} \rightarrow \mathcal{A}) \rightarrow \mathcal{B}))$$

con la sola regola del *modus ponens.*

generi» assumendo per k – come noi faremo – valori finiti maggiori o *uguali* a zero); i connettivi proposizionali ($\neg$, $\wedge$ $\vee$, $\rightarrow$, $\leftrightarrow$) e i quantificatori ($\exists$, $\forall$).

Le formule saranno definite ancora induttivamente (si noti la condizione *a*) sulle atomiche): *a*) se P_i^k è una lettera predicativa a k posti e $x_1, x_2, ..., x_k$ sono variabili, allora $P_i^k(x_1, x_2, ..., x_k)$ è una formula (formule atomiche); se $\mathscr{A}$ è una formula $(\neg\mathscr{A})$ è una formula; se $\mathscr{A}$ e $\mathscr{B}$ sono formule anche $(\mathscr{A} \wedge \mathscr{B})$, $(\mathscr{A} \vee \mathscr{B})$, $(\mathscr{A} \rightarrow \mathscr{B})$, $(\mathscr{A} \leftrightarrow) \mathscr{B})$ sono formule; se $\mathscr{A}$ è una formula e x una variabile, allora anche $(\forall x \mathscr{A})$ e $(\exists x \mathscr{A})$ sono formule; *c*) nient'altro è una formula.

Gli assiomi del calcolo predicativo sono ora:

1) gli assiomi enunciativi (opportunamente formulati nel nuovo linguaggio);

2) $\forall x P(x) \rightarrow P(y)$

3) $P(y) \rightarrow \exists x P(x)$

(ove abbiamo adottato evidenti abbreviazioni); oltre alla regola di separazione si aggiungono altre due regole che possono indicarsi compendiosamente con i seguenti schemi

$$\frac{\mathscr{A} \rightarrow \mathscr{B}(x)}{\mathscr{A} \rightarrow \forall x\, \mathscr{B}(x)} \quad \text{e} \quad \frac{\mathscr{B}(x) \rightarrow \mathscr{A}}{\exists x\, \mathscr{B}(x) \rightarrow \mathscr{A}}$$

purché in entrambi gli schemi x non figuri libera in $\mathscr{A}$. Occorrerà, naturalmente, una regola di sostituzione per variabili individuali libere e una per variabili predicative.[29]

[29] Ricollegandoci alla nota precedente, il sistema si può estendere ad esempio assumendo:

1) Gli assiomi enunciativi
2) Gli schemi di assiomi

$$\forall x\, \mathscr{A}(x) \rightarrow \mathscr{A}(y); \ \forall x(\mathscr{A} \rightarrow \mathscr{B}) \rightarrow (\mathscr{A} \rightarrow \forall x\, \mathscr{B})$$

(con opportune limitazioni sulle variabili y e x) e, oltre alla regola di separazione, la sola regola di generalizzazione $\mathscr{A} / \forall x \mathscr{A}$, e una di sostituzione per variabili individuali libere. Nel seguito ci riferiremo sempre, all'occorrenza, a questo sistemi «più semplici» ribadendo tuttavia la loro equivalenza espressiva e deduttiva col sistema di Hilbert-Bernays qui presentato. A questo proposito va ancora aggiunto che al primo ordine si conviene in generale di considerare l'identità come un

La definizione di teorema si estende nel modo ovvio e conserva il suo carattere induttivo.

Lungo linee analoghe si può procedere nel caso della logica del secondo ordine e di ordine superiore. Fatto questo, è naturale chiedersi come, all'interno di un simile quadro di riferimento, si possa procedere alla formalizzazione di teorie specifiche.

Ovviamente l'esplicitazione formalizzata di una teoria in un dato linguaggio comprenderà, oltre che gli assiomi logici, anche assiomi specifici per la teoria, ossia delle proposizioni formalizzate la cui interpretazione è destinata per così dire a limitare le possibili realizzazioni della teoria stessa a quelle che appunto ne soddisfano gli assiomi. Assumiamo per ora questi termini in senso intuitivo (con perfetta aderenza storica del resto, ché infatti la precisazione rigorosa dell'aspetto semantico interpretativo non sarà un risultato acquisito che negli anni trenta). Prendendo come esempio l'aritmetica avremo un sistema elementare del tipo seguente (che indicheremo come sistema $\mathfrak{P}$, da Peano).

1) Un qualunque sistema di assiomi per il calcolo dei predicati con identità.

2) Assiomi specifici

2.1 $\forall x(\neg s(x) = \overline{0})$

2.2 $\forall x \forall y(s(x) = s(y) \rightarrow x = y)$

2.3 $\mathscr{A}(\overline{0}) \wedge \forall x(\mathscr{A}(x) \rightarrow \mathscr{A}(s(x))) \rightarrow \forall x \mathscr{A}(x)$

2.4 $\forall x \forall y(x+\overline{0} = x \wedge x+s(y) = s(x+y))$

2.5 $\forall x \forall y(x \cdot \overline{0} = \overline{0} \wedge x \cdot s(y) = ((x \cdot y) + x))$[30]

predicato costante il cui comportamento è regolato da opportuni assiomi (di solito 2). È quindi ormai uso comune inserire tale predicato (e gli assiomi relativi) nella parte *logica* di una teoria: si parla allora, in questo caso, di teorie con identità o di *teorie elementari*.

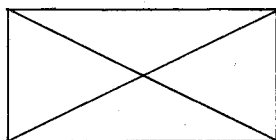
[30] 2. Ovviamente facciamo l'ipotesi di aver stabilito il linguaggio in modo tale da poter esprimere, e in modo sensato, ossia come formule, gli assiomi precedenti. Intuitivamente, «$s(x)$» va letto come «successore di x», gli altri *segni* sono intesi nel loro significato usuale. 2.1 afferma allora che 0 non è successore di alcun numero; 2.2 che se due numeri hanno successori uguali sono uguali; 2.4 e 2.5 definiscono induttivamente le operazioni di somma e prodotto; 2.3, infine, esprime l'assioma di induzione ossia afferma che se 0 gode della proprietà espressa da una certa formula $\mathscr{A}$ e se per un qualunque numero, se esso gode di quella proprietà anche il suo successore ne gode, allora tutti i numeri godono di quella proprietà. Si noti che 2.3 non è un assioma bensì uno *schema* (metalinguistico) di as-

Come osservato in nota 30, utilizzando il linguaggio elementare, non siamo in grado di formulare l'assioma di induzione appunto come assioma ma solo come schema. La situazione cambia drasticamente se utilizziamo il linguaggio del secondo ordine, e in questo caso, utilizzando la quantificazione delle proprietà, 2.3 diviene una vera e propria formula. Una situazione analoga si presenta in altri casi, formalizzando teorie matematiche significative, ad esempio – come sottolineano Hilbert e Ackermann – per la geometria elementare assiomatizzata da Hilbert, in cui il principio di completezza esige un linguaggio del secondo ordine. Hilbert ed Ackermann concludono così che la logica del primo ordine non è sufficiente per formulare tutte le teorie matematiche significative e mentre ammettono che essa basta per dedurre singolarmente le conseguenze delle diverse teorie, affermano che: «Appena l'oggetto di investigazione diviene la fondazione di... teorie matematiche, appena uno desidera determinare in quale relazione la teoria sta con la logica ed entro che limiti si può ottenerla da concetti e operazioni puramente logici, essenziale diviene la logica del secondo ordine». Vedremo più avanti come la questione dei rapporti fra primo e secondo ordine sia inestricabilmente connessa con i fondamenti della teoria degli insiemi e col rapporto fra logica e matematica.

Vogliamo ora occuparci di un altro problema, che è connesso con il frammentarsi della Grande Logica e l'emergere di alternative sul piano logico. Abbiamo già parlato delle ricerche di Post concernenti i sistemi a più valori. Se Post non concepiva le proprie ricerche come contrapposte all'impostazione dei *Principia*, ben diverso è l'atteggiamento che, sin dal 1912, Lewis assunse nei confronti della logica sottogiacente ai *Principia* e soprattutto della implicazione materiale. Fu proprio dalla ricerca di un punto di contatto fra analisi metamatematica di concetti quali coerenza, deducibilità, ecc. e logica proposizionale che Lewis fu portato ad introdurre la sua nozione di implicazione stretta basandosi su

siomi, dal quale cioè si possono ricavare infiniti assiomi sostituendo ad $\mathscr{A}$ (variabile metalinguistica) una qualunque formula con una variabile libera del linguaggio. 2.3 non esprime completamente il contenuto del principio di induzione; per farlo dovrebbe essere esteso a *tutte* le proprietà (e non solo a quelle esprimibili nel nostro linguaggio) ma sappiamo che per poter esprimere ciò occorrerebbe passare al secondo ordine. Riprenderemo più avanti la questione.

distinzioni modali. Come si ricorderà già McColl aveva intrapreso un simile tentativo ma è Lewis, prima da solo nel *Survey* già citato, poi nel 1932 in collaborazione con Cooper Harold Langford in *Symbolic Logic*, che getta le basi di una vera e propria logica modale simbolica, ricollegandosi a tutta la tradizione intensionale che il recente sviluppo della logica sembrava aver rigettato.

Prima dell'inizio dell'era moderna non era mai venuta meno la convinzione che lo studio delle nozioni modali – necessario, possibile, contingente, impossibile – rientrasse a pieno diritto nell'ambito della logica. Anzi, la prima trattazione sistematica delle modalità si ritrova, accanto alla logica degli enunciati categorici, nell'*Organon* aristotelico. Va osservato tuttavia che nell'opera di Aristotele, in cui viene sviluppata una teoria notevolmente matura e articolata del sillogismo modale, manca una definizione rigorosa e univocamente determinata delle nozioni modali fondamentali. Negli *Analitici* infatti Aristotele sembra intendere il possibile come «ciò che non è né necessariamente falso né necessariamente vero», mentre nel *De interpretatione* propende per la definizione del possibile come «ciò che non è necessariamente falso». Questa seconda accezione è quella ereditata da Teofrasto, suo successore alla direzione del Liceo, e con lui da tutta la logica successiva. A partire da Teofrasto i rapporti di interrelazione tra le modalità si possono esprimere mediante un quadrato di opposizioni così costruito:

e mediante le definizioni: $\Box p = \neg \Diamond \neg p$ e $\Diamond p = \neg \Box \neg p$.

[31] La definizione qui adottata del contingente è più comprensiva della definizione, pure di uso corrente, secondo cui una proposizione è contingente quando per stabilire la sua verità o falsità è decisiva l'esperienza (e pertanto quando non è né necessaria né impossibile: $\neg \Box p \wedge \neg \Diamond p$). Nella logica antica e medievale la nozione era ancora più equivoca: molto spesso il contingente veniva identificato con il possibile, o viceversa. Secondo Bochenski questa confusione si ritrova anche in Teofrasto.

Ma già la corretta «lettura» degli enunciati modali presenta dei problemi. Per usare la terminologia dei medioevali, le modalità si possono intendere come modalità *de dicto* o *sensu composito* oppure come modalità *de re* o *sensu diviso*: nel primo caso la modalità viene intesa come attributo di un *enunciato* (ad esempio «"Socrate corre" è possibile») mentre nel secondo caso è intesa come *modo della predicazione*, ossia dell'attribuzione di un predicato al soggetto («Socrate è possibile che corra»). Secondo Oskar Becker e Innocenzo Bochenski, Aristotele intendeva le modalità come modalità *de re*. La questione del rapporto fra i due tipi di modalità non era comunque oggetto di discussione tanto nell'antichità quanto nel medioevo cristiano, in cui continuava ad esser viva e operante la distinzione tra enunciati modali ed enunciati *de inesse* o *de puro inesse* (ossia categorici). Analisi ragguardevoli delle modalità si ritrovano in Pietro Ispano, Abelardo, Alberto Magno e nell'opuscolo *De modalibus propositionibus* attribuito a S. Tommaso. La logica araba del medioevo è pure interessata alle modalità ma con sfumature diverse, che derivano dall'eredità dei megarici e degli stoici e dalla loro tendenza a legare le modalità allo studio del tempo (in questo atteggiamento interviene anche l'interesse per i problemi del fatalismo e della predestinazione: per il fatalista ciò che accade accade *necessariamente*, e cioè $p \rightarrow \Box p$).

Questa tradizione si esaurisce con l'era moderna. Le modalità cessano di essere di competenza della logica e passano nell'area della filosofia: i *modi* della predicazione (così sono ormai considerate le modalità) non riguardano tanto le proposizioni quanto i giudizi, considerati come gli *atti mentali* che vengono espressi nella proposizione. Questo modo di intendere le modalità trova la sua espressione più caratteristica nell'*Analitica trascendentale* kantiana, in cui le modalità compaiono come categorie dell'intelletto dedotte appunto dalla tavola dei giudizi.

La circostanza più interessante dal punto di vista storico non è però l'abbandono della tradizione logico-modale nell'era moderna, quanto piuttosto il fatto che la logica formale del XIX secolo e gli stessi *Principia mathematica* nascono ignorando completamente le modalità e persino il problema di una loro eventuale collocazione nell'ambito della logica. Se si ricorda la polemica tra McColl e Russell, si potrebbe addirittura dire che l'obiettivo di Russell fu quello di eliminare decisamente, con la teoria dell'implicazione

formale, ogni riferimento a nozioni modali. E la cosa può risultare tanto più sorprendente quando si pensi che non mancavano gli strumenti, in particolare a partire dall'algebra delle classi di Boole (e quindi di Schröder) per la fondazione di una logica delle modalità. Ciò dipendeva, nel caso ad esempio della sistemazione logicista, e dall'aver ammesso una rigida semantica, per cui ogni proposizione, estensionalmente, può interpretarsi solo come vera o come falsa (e non ad esempio come talora vera e talora falsa ossia come possibile) e nell'aver introdotto su questa base i connettivi logici, in particolare l'implicazione materiale, come funzioni di verità considerate estensionalmente.

Non stupisce perciò che non solo in McColl e Lewis, ma anche in altri ricercatori come Henry Bradford Smith, che lavorò negli anni trenta in una prospettiva analoga, esista un legame profondo fra algebra della logica, rigetto dell'implicazione materiale e attenzione ai problemi metamatematici. Il bersaglio polemico diretto di Lewis non era infatti tanto il «dogma» della bivalenza, quanto l'implicazione materiale di Russell come base dell'inferenza; egli infatti si propone di «sviluppare un calcolo proposizionale che non è ristretto a relazioni estensionali o materiali... ed è basato su una relazione di implicazione che non ha queste proprietà peculiari». Le «proprietà peculiari» alle quali allude Lewis sono i cosiddetti paradossi dell'implicazione materiale cui sopra abbiamo accennato che si ricavano direttamente dalla sua stessa definizione. Essi possono esprimersi dicendo che una proposizione vera è implicata da qualunque proposizione $[p \rightarrow (q \rightarrow p)]$ o che una proposizione falsa implica qualunque proposizione $[\neg p \rightarrow (p \rightarrow q)]$ o, ancora, che di due qualunque enunciati p e q uno implica sempre l'altro o viceversa. Fatto ancora più grave, in vista degli interessi di Lewis se si definisce «p è coerente con q» come «p non implica la falsità di q» $[\neg (p \rightarrow \neg q)]$ e «q è indipendente da p» come «p non implica q» $[\neg (p \rightarrow q)]$, allora in termini di implicazione materiale due (o un numero finito di) proposizioni qualunque (in particolare: un sistema di assiomi per il calcolo proposizionale) non possono mai essere contemporaneamente coerenti e indipendenti. (Si noti che ciò significa che se si interpretasse la relazione di implicazione materiale come relazione di deducibilità, non avrebbe senso richiedere contemporaneamente l'indipendenza e la coerenza degli assiomi.) Ora, secondo Lewis,

ciò deriva dall'aver limitato lo studio della logica al momento estensionale trascurando le relazioni intensionali (ossia di contenuto) fra le proposizioni, sicché «è particolarmente desiderabile che la logica delle proposizioni sia sviluppata in modo tale che venga incluso il significato usuale di "implica", che è intensionale». Egli vuol quindi sviluppare un calcolo basato su un significato di «implica» in modo tale che «"p implica q" sia sinonimo con "q è deducibile da p"», dove però – si badi – «deducibile» non significa «deducibile entro uno specifico sistema, in particolare quello modale che si costruisce», ma «deducibile in base a qualche sistema inferenziale corretto per quell'applicazione». L'obiettivo è quindi una logica che – ammettendo i legami intensionali – consente diversi tipi di schemi di inferenza e diverse classificazioni delle proposizioni e accetta esplicitamente la molteplicità dei criteri che applichiamo nel classificare gli argomenti.

Solo a questo patto per Lewis è possibile sperare che la logica permetta l'analisi di proprietà metamatematiche fondamentali quali l'indipendenza, la coerenza, ecc. L'obiettivo è una logica che inglobi in sé quegli aspetti *meta*teorici che per Lewis non sono affrontabili con gli strumenti dei *Principia*.

Ora secondo Lewis q è deducibile da p non quando «$p \rightarrow q$» è vera bensì quando essa è *necessariamente* vera ossia quando è *impossibile* che sia falsa. Il sistema della deducibilità introdotto da Lewis era quindi destinato a essere anche un sistema delle modalità *logiche*. In esso, in luogo dell'implicazione materiale compare la relazione di deducibilità sotto forma di una «implicazione stretta» rappresentata dal simbolo «$\prec$» e che in termini di implicazione materiale può definirsi come segue

$$p \prec q = \neg \Diamond (p \wedge \neg q) \qquad \text{ossia} \qquad \Box (p \rightarrow q)$$

(dove il segno $\Diamond$ sta per «possibile» e il segno $\Box$ per «necessario»). La teoria dell'implicazione materiale si ricava come sottosistema del sistema dell'implicazione stretta, senza che sia possibile identificare le due relazioni in quanto mentre è possibile dimostrare $(p \prec q) \prec (p \rightarrow q)$ non è invece derivabile come teorema l'affermazione inversa $(p \rightarrow q) \prec (p \prec q)$. Lo stesso vale per gli altri quattro sistemi, via via deduttivamente più forti, che Lewis costruisce sulla base del primo aggiungendo come assiomi enunciati

che, per quanto plausibili intuitivamente, non sono derivabili nel sistema iniziale. Se, con Lewis, identifichiamo con **S1** il sistema di base in cui vengono canonizzati «gli stretti principi dell'inferenza deduttiva», gli altri sistemi (**S2**-**S5**) si presentano come specificazioni ottenute dal sistema originale mediante la postulazione di ulteriori proprietà della relazione d'implicazione. Queste proprietà non risultano dalla pura nozione di inferenza deduttiva, ma riflettono situazioni generali che possono presentarsi in casi specifici.

Abbiamo così la tavola seguente, in cui si dà il sistema **S1** e vengono schematizzati i rapporti tra i vari sistemi:
Assiomi di **S1**

1. $(p \wedge q) \prec (q \wedge p)$
2. $(p \wedge q) \prec p$
3. $p \prec (p \wedge p)$
4. $((p \wedge q) \wedge r) \prec (p \wedge (q \wedge r))$
5. $((p \prec q) \wedge (q \prec r)) \prec (p \prec r)$
6. $(p \wedge (p \prec q)) \prec q$[32]

Naturalmente al sistema sono associate delle regole che possono ad esempio essere: una regola di sostituzione (del tutto analoga a quella data nel caso del calcolo classico); una regola di sostituzione di equivalenti stretti (sostituendo a una sottoformula di un teorema una formula ad essa strettamente equivalente si ottiene ancora un teorema); una regola di separazione del tutto analoga a quella classica (riferita però qui, ovviamente, all'implicazione stretta). Si ha allora

$$\mathbf{S1} \rightarrow \mathbf{S2}\ (= \mathbf{S1} + \Diamond (p \wedge q) \prec \Diamond p)$$
$$\mathbf{S2} \rightarrow \mathbf{S3}\ (= \mathbf{S1} + (p \prec q) \prec (\Diamond p \prec \Diamond q))$$
$$\mathbf{S3} \rightarrow \mathbf{S4}\ (= \mathbf{S1} + \Diamond \Diamond p \prec \Diamond p)$$
$$\mathbf{S4} \rightarrow \mathbf{S5}\ (= \mathbf{S1} + \Diamond p \prec \Box \Diamond p)$$

dove la freccia indica che il sistema a destra è deduttivamente più forte di quello a sinistra e il simbolo « + » sta a indicare che il si-

[32] Gli altri connettivi devono pensarsi introdotti per definizione a partire da $\neg$, $\wedge$ e $\Diamond$ come segue $(\alpha \wedge \beta) = \neg(\neg \alpha \wedge \neg \beta)$; $(\alpha \prec \beta) = \neg \Diamond (\alpha \wedge \neg \beta)$. $(\alpha \equiv \beta) = (\alpha \prec \beta) \wedge (\beta \prec \alpha)$; $\Box \alpha = \neg \Diamond \neg \alpha$.

stema in questione si ottiene da **S**1 con l'aggiunta degli assiomi indicati.

Fatto importante, i sistemi sono distinti e gli assiomi «supplementari» risultano quindi indipendenti dal sistema base, cosicché rappresentano, come nell'intenzione, genuine specificazioni ulteriori delle proprietà della relazione di implicazione; che tipo di specificazioni in particolare non è immediato vedere considerando solamente gli assiomi. Il loro significato semantico non è chiaro intuitivamente e questo spiega come Lewis (e il suo allievo William Parry cui si deve la prima analisi dettagliata dei mutui rapporti fra i sistemi) fosse costretto, per ottenere le menzionate dimostrazioni di indipendenza, a ricorrere all'uso di «tavole di verità» con più di due valori di verità (matrici) per i singoli connettivi, costruite *ad hoc*, caso per caso. L'esigenza dei singoli assiomi aggiuntivi era così frutto, più che di considerazioni semantiche sistematiche, di analisi di carattere tutto sommato «sperimentale».

Il problema di un'analisi semantica completa dei sistemi modali **S**1-**S**5 rimarrà l'ostacolo più grave alla diffusione delle idee di Lewis e la ragione principale di una loro «subordinazione» alla nascente logica polivalente. Ancora la logica modale è infatti la motivazione diretta e più immediata da cui prende le mosse Jan Łukasiewicz (1878-1956) per l'introduzione delle *logiche polivalenti*, la cui prima apparizione risale a una brevissima memoria (una pagina) da lui presentata nel 1920, *Sulla logica trivalente*. In questa occasione egli motiva la cosa in termini generali, in quanto ritiene che la logica trivalente che ivi presenta abbia soprattutto importanza teoretica come invito a costruire un sistema di logica non aristotelica (più tardi preferirà suggerire il nome *non crisippea*). «Se questo nuovo sistema di logica abbia importanza pratica» prosegue «si vedrà solo quando i fenomeni logici, in particolare quelli delle scienze deduttive, saranno esaminati più a fondo e quando le conseguenze della filosofia indeterministica, che è il substrato scientifico della nuova logica, potranno essere confrontate con i dati empirici.» Successivamente restringe queste motivazioni in termini più precisi, affermando ad esempio che «il sistema di logica trivalente deve la sua origine a certe ricerche da me condotte sulle cosiddette proposizioni modali e sulle nozioni di possibilità e necessità strettamente connesse ad esse», fino ad af-

fermare decisamente: «...come fondatore dei sistemi di logica proposizionale polivalenti affermo che teoricamente questi sistemi non sono stati sviluppati sulla base del convenzionalismo o relativismo, ma sono emersi da ricerche logiche relative alle proposizioni modali e ai concetti annessi di possibilità e necessità».[33]

Łukasiewicz prende quindi le mosse dalla considerazione delle proposizioni modali: «È possibile che *p*», che egli simbolizza con M*p*, «Non è possibile che *p*» (NM*p*), «È possibile che non *p*» (MN*p*) e «Non è possibile che non *p*» (ossia è necessario che *p*, NMN*p*).[34] Ora egli nota tre proposizioni modali direttamente derivate da principi classici della modalità[35] altamente plausibili da un punto di vista intuitivo, e che tuttavia portano, se interpretate classicamente, a risultati «spiacevoli» o addirittura a contraddizioni. Tali proposizioni sono:

I. Se non è possibile che *p*, allora non *p*: $\neg \Diamond p \rightarrow \neg p$ [CNM*p*N*p*]

II. Se si suppone che non *p*, allora (sotto questa ipotesi) non è possibile che *p*: $\neg p \rightarrow \neg \Diamond p$ [CN*p*NM*p*]

III. Per qualche *p*: è possibile che *p* ed è possibile che non *p*: $\exists p (\Diamond p \wedge \Diamond \neg p)$ [Σ*p*KM*p*MN*p*].

[33] Si ricordi in proposito quanto detto su Post circa la sua presentazione delle logiche polivalenti. Dirà Łukasiewicz, attorno al 1930: «È vero che Post ha indagato i sistemi proposizionali polivalenti da un punto di vista puramente formale, tuttavia non gli è riuscito di interpretarli logicamente. I ben noti tentativi di Brouwer che rigetta la validità universale della legge del terzo escluso e quindi rifiuta molte tesi dell'ordinario calcolo proposizionale non hanno finora condotto a un *sistema* basato intuivamente.» Per quest'ultima considerazione si osservi che, al tempo in cui Łukasiewicz scriveva la sua nota non era stata ancora pubblicata la «formalizzazione» di Heyting della logica e della matematica intuizioniste. Nel 1952 in *On the intuitionistic theory of deduction* (*Sulla teoria intuizionista della deduzione*) dirà invece: «Mi sembra che fra i sistemi di logica polivalente finora noti la teoria intuizionista sia il più intuitivo ed elegante.»

[34] Scritture di questo tipo intervengono nell'ambito della cosiddetta «notazione polacca» introdotta appunto da Łukasiewicz, che permetteva di evitare l'uso delle parentesi nelle formule semplicemente premettendo gli operatori proposizionali ai loro argomenti: così invece di scrivere $\neg \alpha$, $\alpha \wedge \beta$, $\alpha \rightarrow \beta$, $\alpha \leftrightarrow \beta$, Łukasiewicz scriveva rispettivamente $N\alpha$, $K\alpha\beta$, $A\alpha\beta$, $C\alpha\beta$, $E\alpha\beta$. E ad esempio una formula quale $(\alpha \rightarrow \neg \beta) \rightarrow (\neg \gamma \leftrightarrow \delta)$ viene resa in notazione polacca come $CC\alpha N\beta EN\gamma\delta$. Per comodità del lettore noi renderemo le formule nella nostra abituale notazione, indicando semmai *anche* la notazione polacca.

[35] Ad esempio, *Ab oportere ad esse valet consequentia*; *A esse ad posse valet consequentia*; e, per contrapposizione da quest'ultima, *Ab non posse ad non esse valet consequentia*.

I teoremi II e III hanno conseguenze «non piacevoli»; ad esempio dal secondo si può ricavare «se è possibile p allora p» (se è possibile che il paziente muoia allora muore), dal III si può derivare un teorema secondo il quale ogni proposizione è possibile (con risultati del tipo «è possibile che 2 sia primo» e «è possibile che 2 non sia primo»). Ma si giunge addirittura a contraddizione considerando congiuntamente le proposizioni II e III. Infatti, utilizzando le conseguenze prima dette dei due teoremi si ha per separazione

$$\frac{\begin{array}{l}\Diamond p \rightarrow p \\ \Diamond p\end{array}}{p}$$

ossia si ottiene semplicemente p; ogni proposizione risulterebbe così essere un teorema il che, come sappiamo, è un possibile modo per esprimere la contraddittorietà di un sistema proposizionale.

Ora queste leggi conducono a contraddizione solo se sono interpretate in un sistema di logica classica se si assume la bivalenza delle proposizioni, ove cioè si assuma che una proposizione possa avere come valore soltanto il Vero o il Falso. È proprio per evitare questa banalizzazione delle modalità che per Łukasiewicz è essenziale introdurre una logica trivalente; sulla base della definizione $\Diamond p = \neg p \rightarrow p$ [Mp = CNpp] (dovuta all'allora suo allievo Alfred Tarski) è così possibile far vedere che «tutti i tradizionali teoremi della logica proposizionale modale possono essere stabiliti senza contraddizione nel calcolo proposizionale trivalente».

Va notato però che in questo modo, differentemente da Lewis e dai suoi sistemi **S**1-**S**5, la logica modale veniva a presentarsi come un *sottosistema* di quella classica e quindi non come una sua generalizzazione. Se così Łukasiewicz riusciva a dare un'analisi semantica precisa della nozione di sistema modale, era solo a patto di una rinuncia; proprio quella rinuncia che Lewis non era disposto a fare e che lo spingeva a ritenere che l'approccio polivalente di Łukasiewicz non fosse in grado di fornire una risposta soddisfacente al problema della caratterizzazione generale della nozione di inferenza deduttiva.

Va tuttavia considerato il fatto che – come si è visto – l'orizzonte nel quale si pone Łukasiewicz è diverso e in un certo sen-

so più ideologicamente connotato di quello di Lewis, legato com'è al problema del determinismo e del libero arbitrio e la costruzione di queste nuove logiche polivalenti gli sembra offrire una soluzione a quella «costrizione razionale» cui è sottoposta la scienza fondata «sulla logica aristotelica». «Questa nuova logica», afferma infatti Łukasiewicz, «introducendo il concetto di possibilità oggettiva, distrugge il vecchio concetto di scienza basata sulla necessità. I fenomeni possibili non hanno cause, per quanto essi stessi possano essere l'inizio di una catena causale. L'atto di un individuo creativo può essere libero e al tempo stesso influenzare il corso del mondo. La possibilità di costruire sistemi logici diversi mostra che la logica non è ristretta alla riproduzione dei fatti ma è un libero prodotto dell'uomo come un'opera d'arte. La coercizione logica si dissolve proprio alla sua sorgente». È un atteggiamento che richiama, nelle sue venature spiritualiste, gli analoghi tentativi più o meno contemporanei di Vasil'ev di sviluppare una logica polivalente come logica della possibilità.

A parte comunque questo riferimento ad una problematica filosofica «tradizionale», Łukasiewicz e tutta una schiera di collaboratori – fra i quali vanno ricordati almeno Adolf Lindenbaum (1904-1941), Alfred Tarski e Mordechaj Waisberg – fra il 1920 e il 1930 generalizza e sistematizza i risultati ottenuti sulle logiche polivalenti e più in generale proposizionali nel già citato articolo del 1930, a firma di Tarski e Łukasiewicz, dal titolo *Ricerche sul calcolo proposizionale*. Qui vengono considerate logiche L_n a un numero finito qualunque o infinito numerabile di valori (ossia per n uguale a un numero naturale qualunque o $n = \aleph_0$), dando formule generali per le tavole di verità dei connettivi di implicazione e negazione.

Al di là dei risultati relativi alle logiche polivalenti (sui quali ritorneremo nel cap. VI, par. 5) quello che è caratteristico ed estremamente significativo in questo articolo è l'impostazione generale di tipo *metalogico*, che riflette l'indirizzo di ricerche di un decennio della scuola polacca[36] orientate verso lo studio *globale* dei sistemi

[36] Si dovrebbe in realtà parlare di *due* scuole logiche polacche, con caratteristiche e tendenze assai diverse fra loro, e che originariamente fanno capo rispettivamente all'università di Cracovia e a quella di Leopoli (le uniche università polac-

logici e delle teorie e che ottenne alcuni dei suoi risultati più significativi proprio nello studio «della più semplice disciplina deduttiva, il calcolo proposizionale».

Questa tendenza alla considerazione metamatematica delle teorie aveva in Łukasiewicz motivi filosofici precisi ma fu sviluppata con particolare vigore da Tarski, formatosi in un mondo matematico in cui campeggiavano figure come ad esempio Stefan Banach (1892-1945), Kasimierz Kuratowski e Waclaw Sierpiński (1882-1969); si concretizza così quell'idea di una «metascienza» generalizzata, non più limitata alla logica proposizionale ma estesa a *tutte* le teorie, matematiche e non, fondata sui metodi insiemistici infi-

che sotto la dominazione asburgica). La matrice della scuola di Cracovia è nettamente scientifica (matematica) e sembra doversi far risalire all'attività di Jan Ślezsyński e quindi del fisico Stanislaw Zaremba, dei quali fu allievo quel Leon Chwistek di cui si è già parlato, a proposito della sua proposta di «revisione» della teoria ramificata dei *Principia*, con la sostituzione ad essa della teoria semplice dei tipi. Allievi di Chwistek furono W. Hepter e J. Herzberg (massacrati, come molti altri studiosi ebrei polacchi dell'epoca, dai nazisti) che diedero a loro volta notevolissimi contributi a quello che John Myhill chiama «il più importante tentativo mai compiuto per stabilire la non contraddittorietà della matematica». La scuola di Leopoli (Lvov) viceversa (che verrà poi detta di Leopoli-Varsavia perché alcuni dei suoi maggiori rappresentanti si trasferiranno appunto a Varsavia) ha matrice nettamente filosofica e può farsi risalire all'attività di Kasimierz Twardowski (1886-1938) il quale, pur non essendo professionalmente un logico, a partire dal 1895 preparò col suo insegnamento «la base per lo sviluppo futuro in senso logico, abituando i suoi numerosi allievi all'uso del metodo scientifico in filosofia, instillando nelle loro menti un grande rispetto per il pensiero chiaro e per la precisione espressiva e in generale fondendo il pensiero filosofico coi risultati di un abito intellettuale scientifico» (Jordan). Dal 1906 si affianca a Twardowski il giovane allievo Łukasiewicz che già nel 1910 pubblica il volume *Il principio di non contraddizione in Aristotele*, impostato secondo lo stile delle moderne ricerche di logica, anche se il vero e proprio inizio delle ricerche sistematiche in questa direzione va posto, secondo Zbigniew A. Jordan, negli anni immediatamente successivi alla prima guerra mondiale e abbia come sua prima fase uno studio approfondito dei *Principia mathematica*: è nel 1927 che Zygmund Zawirski (1882-1948) pubblica in *Il rapporto fra logica e matematica alla luce delle ricerche moderne* un riassunto del manifesto logicista. Fra gli allievi di Łukasiewicz della prima generazione vanno ricordati almeno Tadeusz Kotarbiński (1886-1981) e Kasimierz Ajdukiewicz (1890-1963). Nel 1912 giunge a Leopoli Stanislaw Leśniewski (1886-1939) che qui si accosta alla problematica della logica moderna e cui avremo occasione di accennare più avanti per i suoi lavori sui fondamenti della teoria degli insiemi. Nel 1915 Łukasiewicz si trasferisce a Varsavia dove viene raggiunto, dopo la seconda guerra mondiale, da Leśniewski. I due diventano gli ispiratori indiscussi della scuola logica di Varsavia, alla quale appartengono appunto fra gli altri, Tarski (che diventa docente nel 1926) e quindi Lindenbaum, Wajsberg, M. Presburger ecc.

nitari della topologia e dell'algebra astratta di cui già abbiamo parlato nel cap. II.

3.2 *Le critiche al sistema di Zermelo. Proposte alternative*

Nel paragrafo 2.2 abbiamo presentato il sistema assiomatico con cui Zermelo riteneva di poter ricostruire la teoria cantoriana degli insiemi evitando nel contempo le antinomie e avevamo anche accennato alle considerazioni di Zermelo sulla coerenza e sull'indipendenza del proprio sistema di assiomi. Vogliamo ora riferire sui risultati del complesso e assiduo lavoro di rielaborazione critica del sistema zermeliano che ha inizio proprio nel decennio che stiamo qui considerando e che costituisce l'inizio di quello studio sistematico della teoria *assiomatica* degli insiemi che prendeva il posto della teoria «ingenua» (*naive*) cantoriana.

Va detto subito che nulla di più di quanto affermato da Zermelo fu di fatto ottenuto in questo periodo relativamente alla pura e semplice *coerenza* del sistema. Anche nell'ambito delle rinnovate riflessioni dei vari autori di questo periodo non si ricostruiscono cioè nel sistema le note antinomie, ma non si guadagna in alcun senso un qualche contributo alla formale dimostrazione della sua coerenza.

Del resto era chiaro per tutti che in quanto teoria fondante, in cui interpretare tutte le altre teorie matematiche, non era possibile alcuna dimostrazione di coerenza *relativa* per la teoria degli insiemi, come quelle fornite da Hilbert per i vari sistemi geometrici. Il problema diveniva così quello di trovare una dimostrazione *assoluta* di coerenza, che non fosse circolare e non presupponesse nozioni insiemistiche. Sarà proprio per giustificare l'uso dell'infinito in matematica, che nella teoria degli insiemi trovava la sua piena realizzazione, che Hilbert giungerà in questi anni a formulare in modo esplicito il suo programma finitista per le dimostrazioni di coerenza.

Per quanto riguarda specificatamente la teoria degli insiemi, l'attenzione si concentrò invece sul problema della categoricità dei diversi sistemi assiomatici, e sullo studio della forma dei possibili modelli, indagando possibili rafforzamenti dell'assioma di isolamento, il ruolo dell'assioma di scelta, e tentando di affiancare alla

teoria degli insiemi una teoria delle classi. Centrale diverrà così l'analisi dell'intergioco fra possibilità di costruzioni insiemistiche e capacità descrittive del linguaggio, portando in primo piano la questione della natura delle proprietà definite e più in generale del linguaggio adeguato alla costruzione della matematica. Saranno questioni di portata più generale, che spesso travalicheranno i limiti della stretta problematica insiemistica, portando ancora una volta ad aspri dibattiti sui rapporti tra logica, linguaggio e matematica.

Punto naturale di partenza possono essere le critiche al sistema di Zermelo che possiamo sostanzialmente suddividere in due tipi:

1) da un lato le lacune di tipo metodologico-filosofico riguardanti il concetto di «proprietà definita» che portano in primo piano la questione del linguaggio in cui formulare la teoria;

2) dall'altro la presenza di «limiti» che si concretizzano in due aspetti tra loro in certo senso opposti: l'impossibilità di assicurare l'esistenza di determinati insiemi che si dimostrano indispensabili per una ricostruzione quanto più possibile completa della teoria di Cantor e l'opportunità di escludere l'esistenza di altri insiemi altamente indesiderabili dal punto di vista intuitivo. In entrambi i casi, comunque, problemi connessi con la categoricità.

Cominciamo dal primo tipo di obiezioni. Il concetto di «proprietà definita», si ricorderà, era stato introdotto da Zermelo per eliminare quelle antinomie che (come vedremo nelle prossime pagine) oggi chiamiamo, seguendo Ramsey, *linguistiche* in contrapposizione a quelle *logiche*, per evitare le quali era sufficiente la «relativizzazione» dell'assioma di comprensione. È chiaro che la stessa vaghezza del concetto non permetteva una esaustiva ispezione di tutte le eventuali proprietà allo scopo di determinare se esse fossero o meno «definite». Si imponeva quindi una precisazione del concetto di «proprietà definita» e tale chiarificazione e precisazione venne intrapresa da Skolem e da Fraenkel indipendentemente l'uno dall'altro nei primissimi anni venti. La soluzione di Fraenkel viene presentata brevemente nell'articolo del 1922 *Der Begriff «definit» und die Unabhängigkeit des Auswahlaxioms* (*Il concetto «definito» e l'indipendenza dell'assioma di scelta*) e più estesamente nell'articolo *Untersuchungen über die Grundlagen der Mengenlehre* (*Ricerche sui fondamenti della teoria degli insiemi*) del 1925. Il me-

todo di Fraenkel, che consiste nel precisare il concetto di definitezza attraverso un particolare impiego della nozione di funzione, è alquanto complicato e non è stato mantenuto nell'uso; il metodo invece che poi è stato universalmente accettato risale a Skolem, che lo presentò in un importantissimo articolo del 1922:[37] *Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre* (*Alcune osservazioni sulla fondazione assiomatica della teoria degli insiemi*). L'idea di Skolem è estremamente semplice e lineare e interviene intrinsecamente, per così dire, nella natura stessa della difficoltà che con l'introduzione del concetto di «proprietà definita» Zermelo voleva evitare. Essa consiste essenzialmente nel precisare il *linguaggio* specifico della teoria degli insiemi in modo

[37] Questo articolo di Skolem – come vedremo – è in effetti estremamente ricco di profonde osservazioni e sulla problematica della teoria degli insiemi in particolare e sul metodo assiomatico in generale. In esso infatti Skolem discute i seguenti 8 punti: 1) istituendo una teoria *assiomatica* degli insiemi Zermelo si fonda sul concetto di *dominio*, il che comporta inevitabilmente una certa circolarità; 2) la questione delle «proprietà definite» di cui si riferisce nel testo; 3) il fatto che in *ogni* assiomatizzazione della teoria degli insiemi le nozioni insiemistiche assumono in modo inevitabile un carattere di *relatività* (Skolem considera questo il punto più interessante del suo lavoro – è infatti la prima enunciazione del cosiddetto «relativismo» di Skolem – e lo fonda sul teorema di Löwenheim del quale, come già detto, parleremo nel paragrafo 4); 4) il fatto che il sistema di Zermelo non è sufficiente ad assicurare una soddisfacente fondazione per la teoria cantoriana, il che porta all'enunciazione dell'assioma di rimpiazzamento; 5) il problema delle definizioni non predicative in rapporto alla questione della dimostrazione di coerenza degli assiomi della teoria degli insiemi; 6) il problema della non univoca determinazione del dominio *B* da parte degli assiomi di Zermelo (anche questo punto è assai importante perché sulla base delle argomentazioni in esso condotte, Skolem avanza in nota la congettura che sia impossibile decidere certe proposizioni, in particolare quella esprimente l'ipotesi (ristretta) del continuo cantoriana); 7) la necessità dell'impiego dell'induzione matematica per «l'indagine logica di sistemi di assiomi dati astrattamente» in particolare relativamente alla loro coerenza (è in questa occasione che Skolem sostiene l'insufficienza delle argomentazioni di Hilbert contro le obiezioni di Poincaré); 8) il senso dell'assunzione del principio di scelta fra gli assiomi, che lo porta a giustificare quei matematici che eventualmente non lo accettano, sulla base del fatto che la maggior parte dei matematici desidera in fin dei conti che la loro scienza «tratti con operazioni di calcolo effettuabili e non consista di operazioni formali su oggetti chiamati così o cosà». Skolem sostiene insomma un atteggiamento costruttivista. Molto interessante è infine la conclusione cui Skolem giunge e cioè che non si debba pensare che il metodo assiomatico abbia quelle caratteristiche di assolutezza che molti sembrano accordargli e in particolare che in nessun caso «gli assiomi della teoria degli insiemi costituiscono la fondazione ideale per la matematica».

induttivo e nell'assumere quindi che proprietà definite siano quelle esprimibili nel linguaggio in esame. Skolem immagina la teoria degli insieme formalizzata in un linguaggio predicativo del primo ordine con identità, le cui costanti (operazioni) logiche sono 1) la congiunzione indicata con giustapposizione, 2) la disgiunzione indicata con $\vee$; 3) la negazione, indicata con $\neg$; 4) la quantificazione universale, indicata con $\forall$; 5) la quantificazione esistenziale indicata con $\exists$,[38] e che ha come costante descrittiva il solo simbolo $\in$ per l'appartenenza. Fatto questo – come scrive Skolem – «per proposizione definita intendiamo ora un'espressione finita costruita da proposizioni elementari della forma $a \in b$ e $a = b$ per mezzo delle cinque operazioni menzionate», in altri termini, una qualunque formula del linguaggio del primo ordine precisato sopra. Ristretto così al linguaggio sopra descritto l'assioma di isolamento cessa di essere un assioma per diventare uno *schema* di assiomi, al pari del principio di induzione.

Malgrado nel suo articolo Skolem non tenti nemmeno di giustificare la propria scelta, limitandosi ad affermare che gli operatori logici adottati sono le «cinque operazioni fondamentali della logica matematica nella notazione di Schröder» non si può negare che dietro questa mossa stava una precisa scelta ideologica, non condivisa da altri protagonisti della ricerca insiemistica, primo fra tutti Zermelo. Perché un linguaggio del primo ordine e non piuttosto un linguaggio del secondo ordine, più ricco ed espressivo? Già abbiamo visto, parlando dei *Grundzüge* di Hilbert e Ackermann, come ancora nel 1928 Hilbert ritenesse essenziale, per sviluppare la matematica, utilizzare linguaggi del secondo ordine o di ordine superiore, proprio in vista della ricerca di categoricità, di completa caratterizzazione degli oggetti matematici sul piano linguistico. Lo stesso Zermelo svilupperà nel 1931 e successivamente nel 1935 idee analoghe, tentando una classificazione in livelli del processo di quantificazione e proponendo addirittura – sulla linea di Schröder – di costruire un vero e proprio linguaggio infinitario con congiunzioni e disgiunzioni infinite; ciò avverrà in due lavori dal titolo *Über Stufen der Quantification und die Logik des Unendlichen* (*Sui livelli di quantificazione e la logica dell'infinito*, 1931) e

[38] Si noti che Skolem impiegava ancora, in questo articolo, la notazione di Schröder.

Grundlagen einer allgmeinen Theorie der mathematischen Satzssysteme (*Fondamenti di una teoria generale dei sistemi di enunciati matematici*, 1935).

L'obiettivo polemico di Zermelo era proprio Skolem. Secondo Zermelo «una sana metamatematica, una vera "logica dell'infinito" diviene possibile solo attraverso una *fondamentale rinuncia* all'assunzione ... che io chiamo "pregiudizio finitista". In ogni caso il vero oggetto della matematica *non* è come molti vorrebbero lo studio delle "combinazioni di segni", bensì le *relazioni ideali da un punto di vista concettuale* tra gli elementi di una *varietà infinita* determinata concettualmente. Così il nostro sistema di segni è sempre uno strumento *incompleto*, oscillante da caso a caso. Esso riflette la nostra comprensione *finita* dell'infinito, che noi non possiamo dominare o comprendere *in modo immediato e intuitivamente*, anche se almeno possiamo avvicinarci ad una qualche padronanza passo per passo».

Avremo occasione di vedere più avanti il risvolto semantico di queste concezioni di Zermelo in rapporto al problema della categoricità; qui vogliamo soffermarci ancora sul problema della scelta del linguaggio in quanto non coinvolgeva solo la teoria degli insiemi, ma la stessa nozione di logica. Ben presto infatti avrebbe preso larga diffusione l'idea che essenzialmente l'oggetto centrale della logica matematica era lo studio dei linguaggi del primo ordine, venendo così a tagliare di netto ogni possibile discorso sul significato di questa scelta. Anche se implicite, c'erano per Skolem delle ragioni per preferire il primo ordine al secondo ordine e alla teoria dei tipi. Come sappiamo, al secondo ordine noi quantifichiamo su tutti i sottoinsiemi di un insieme dato e quindi – a livello interpretativo – assumiamo di poter assegnare un significato determinato all'operazione che da un insieme X ci porta all'insieme dei suoi sottoinsiemi. L'operazione, d'altra parte, è fortemente impredicativa, cosicché l'assunzione di un linguaggio al secondo ordine per descrivere l'universo degli insiemi sembra in qualche modo presupporre, già a livello logico, quello che la teoria deve descrivere, nella fattispecie il concetto di sottoinsieme arbitrario.

È quindi dal punto di vista epistemologico che una fondazione al secondo ordine per la teoria degli insiemi sembra circolare, come del resto – osservava Skolem nell'articolo citato – era la stessa

nozione di *dominio* assunta da Zermelo: non è circolare chiarire il concetto di insieme partendo da quello di dominio? Il risultato centrale dell'articolo di Skolem – come è ricordato nella nota 37 – mostrava appunto la non assolutezza delle nozioni insiemistiche, prima fra tutte di quella di sottoinsieme qualunque, una volta che si assumesse la logica del primo ordine e mostrava così chiaramente come l'assolutezza delle nozioni fosse funzione della logica, creando una circolarità dal punto di vista epistemologico. Sono queste le ragioni profonde che convinsero anche altri matematici, non necessariamente legati ad una concezione antirealista come quella che – vedremo – era propria di Skolem, ad accettare l'idea di formulare gli assiomi della teoria degli insiemi al primo ordine. Sul piano logico questo significava l'emergere della logica elementare come linguaggio fondamentale e non è casuale – in vista di quanto detto sopra – che uno dei primi ricercatori ad isolarlo già ne *Il continuo* del 1918 fosse stato Hermann Weyl sulla cui impostazione predicativista avremo occasione di ritornare.

In ogni caso, quello che risultava indubitabile era che finalmente si disponeva di un sistema ben determinato come base per ogni ulteriore indagine sui limiti e sulle capacità degli assiomi di Zermelo.

Le prime osservazioni al riguardo risalgono a Fraenkel che nell'articolo del 1922 sopra menzionato, sottolinea come in Zermelo non venga isolata in modo adeguato la struttura matematica portante dell'universo degli insiemi. Dice Fraenkel: «...tra le "cose" del "dominio *B*" dalle quali traggono la loro esistenza gli insiemi in virtù degli assiomi, possono trovarsene anche di origine non matematica o addirittura di origine non concettuale. Tale caratteristica è inutile dal punto di vista dell'edificazione della matematica e, almeno da questo punto di vista, va quindi riguardata come insoddisfacente». Veniva cioè criticata l'eccessiva liberalità con cui si riteneva che fosse nello spirito cantoriano l'ammettere «cose qualunque» come entità di partenza dalle quali costruire poi gli insiemi grazie agli assiomi: sono questi gli *Urlemente* o *atomi*, eredi degli individui di tipo zero di russelliana e schröderiana memoria. Fraenkel osservava che l'*unica* caratteristica che da un punto di vista logico va richiesta alle entità fondamentali è quella di essere costituite in modo tale da non ammettere elementi, ossia che nessun'altra entità del dominio sia loro elemento. D'altra parte l'in-

sieme vuoto gode proprio di questa caratteristica e quindi Fraenkel proponeva che l'unica «sostanza» elementare ammessa a livello iniziale fosse esattamente l'insieme vuoto. Ciò veniva ottenuto formalmente enunciando l'assioma di estensionalità, che come si ricorderà Zermelo limitava a insiemi, in modo da valere per *ogni* elemento del dominio. Ne risulta immediatamente che due «cose» qualunque prive di elementi vengono a coincidere fra loro e in particolare ognuna di queste cose coinciderà appunto con l'insieme vuoto.

Più significative dal punto di vista matematico le obiezioni del secondo tipo, dirette contro i limiti di ricostruibilità della teoria di Cantor nel quadro zermeliano. Come dimostrato da Skolem nel citato articolo del 1922, nel sistema di Zermelo non era dimostrabile l'esistenza di insiemi quale il seguente

$$M = \{I, \mathscr{P}(I), \mathscr{P}(\mathscr{P}(I)), \mathscr{P}(\mathscr{P}(\mathscr{P}(I))), \ldots\}$$

il cui primo elemento I è un insieme infinito – la cui esistenza è assicurata dall'assioma dell'infinito – e ogni elemento successivo è l'insieme potenza del precedente. Ognuno degli elementi di M esiste in virtù dell'assioma dell'insieme potenza: quello di cui non si riesce a dimostrare l'esistenza è proprio M nel suo complesso, considerato come insieme (è abbastanza agevole infatti concludere in $\mathfrak{Z}$ dall'esistenza dell'insieme M all'esistenza della classe totale). Questo era particolarmente grave non tanto dal punto di vista della matematica classica, quanto dal punto di vista della teoria cantoriana degli ordinali e dei cardinali. Se infatti il ricorso ad insiemi come quello di sopra, ottenuti collezionando i valori della iterazione di una data operazione su insiemi, non aveva un ruolo centrale nella ricostruzione di aritmetica e Analisi, era essenziale per la teoria cantoniana degli ordinali, in cui si ha a che fare con successioni di ordinali che «devono» avere un limite pena l'impoverimento radicale della teoria medesima. Per rimediare a questa deficienza Skolem proponeva di assumere il cosiddetto *assioma di rimpiazzamento* (o di regolarità o di sostituzione) che da allora è divenuto un'assunzione canonica nella teoria degli insiemi. Tale assioma afferma in sostanza che se si ha un insieme *a* e una *relazione funzionale f* definita su *a* allora esiste un *insieme a'* di tutti e soli i valori di *f* per argomenti tratti da *a*; o in altri termini, se il dominio di una relazione funzionale è un insieme tale è anche il suo codo-

minio. Si noti che la forza dell'assioma sta nel fatto che f è una relazione funzionale (cioè il denotato di una formula $R(x, y)$ del linguaggio per cui si può provare che per ogni $x \in a$ esiste esattamente un y per cui $R(x, y)$) e non solo una funzione, cioè un particolare insieme di coppie appartenente all'universo. Una soluzione analoga a quella di Skolem proponeva nello stesso periodo Fraenkel in almeno tre articoli: *Über die Zermelosche Begründung der Mengenlehre* (*Sulla fondazione zermeliana della teoria degli insiemi*, 1921), *Axiomatische Begründung der transfiniten Kardinalzahlen I* (*Fondazione assiomatica dei numeri cardinali transfiniti, I*, 1922) e nell'articolo del 1922 citato.

L'altra obiezione riguardava – come accennato – il fatto che nella teoria di Zermelo non era esclusa l'esistenza di insiemi *straordinari*, segnalati per la prima volta da Dimitri Mirimanoff nel 1917 in *Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles* (*Le antinomie di Russell e di Burali-Forti e il problema fondamentale della teoria degli insiemi*). Questi insiemi sono caratterizzati dal fatto di contenere una catena discendente infinita o un ciclo di insiemi legati dalla relazione di appartenenza. Se a è uno di tali insiemi si avrà ad esempio che esistono infiniti insiemi $a_1, a_2, \ldots$ tali che

$$\ldots \in a_{n+1} \in a_n \in \ldots \in a_2 \in a_1 \in a.$$

L'indesiderabilità di tali insiemi nasce dal fatto che essi contraddicono l'immagine intuitiva dell'universo degli insiemi che stava dietro il sistema di Zermelo. Si tratta di quella *struttura cumulativa* che in quegli stessi anni von Neumann avrebbe precisato in termini della gerarchia che da lui prende il nome. Sostanzialmente l'idea – fatta l'ipotesi semplificatoria di Fraenkel della non esistenza di *Urelemente* – è che tutti gli insiemi dell'universo si ottengono, a partire dall'insieme vuoto, iterando l'operazione di passaggio all'insieme delle parti. I passi dell'iterazione sono indiciati dagli ordinali appartenenti all'universo, cosicché possiamo definire per ricorsione sugli ordinali (e, si noti bene, la liceità di questa definizione poggia essenzialmente sull'assioma di rimpiazzamento) la gerarchia

$$\begin{aligned} V_0 &= \emptyset \\ V_{\alpha+1} &= \mathscr{P}(V_\alpha) \\ V_\lambda &= \textstyle\bigcup_{\beta<\lambda}(V_\beta), \text{ per } \lambda \text{ ordinale limite.} \end{aligned}$$

Come si può vedere, la gerarchia riflette l'idea intuitiva da cui siamo partiti: si inizia con l'insieme vuoto, agli ordinali successori si passa all'insieme delle parti, e ai passi limite si colleziona quanto ottenuto precedentemente. In questo modo ogni insieme compare per la prima volta in un preciso V_α il cui indice ordinale costituisce il *rango* dell'insieme nella gerarchia. Anche se non in modo diretto come alcuni ritengono, è innegabile che questo modello dell'universo degli insiemi fa da sfondo alla concezione zermeliana per cui ogni insieme può essere isolato solo all'interno di insiemi ottenuti precedentemente e quindi appartiene ad un $V_{\alpha+1}$ dove α è il rango dell'insieme cui si applica l'isolamento. È chiaro a questo punto in che senso gli insiemi straordinari risultino inaccettabili: poiché nella gerarchia ogni insieme ha un rango e dato a se $b \in a$ il rango di b sarà strettamente minore di quello di a, avremo che non può esistere catena discendente infinita e quindi che davvero ogni insieme si ottiene da insiemi di rango minore passando alle parti.

Come garantire però che non esistono insiemi straordinari e che quindi tutti gli insiemi hanno un rango?

Per superare tale difficoltà, dopo vari infruttuosi tentativi di Fraenkel, nel 1928, in *Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre* (*Sulla definizione per induzione transfinita e questioni connesse della teoria generale degli insiemi*) e in *Die Axiomatisierung der Mengenlehre* (*L'assiomatizzazione della teoria degli insiemi*) von Neumann (e dopo di lui, indipendentemente, Zermelo) proponeva di aggiungere un assioma, detto *assioma di fondazione* (o assioma del regresso infinito escluso), il quale postula che ogni insieme non vuoto a contiene un elemento d che non ha con a alcun elemento comune. Si vede facilmente che questo esclude il regresso infinito di cui sopra e comporta quindi, come dimostrerà von Neumann, che l'universo coincide con la gerarchia cumulativa sopra descritta. In questo senso l'assioma – come quello di induzione per i numeri naturali – garantisce una sorta di «forma canonica» dei modelli che ci permette di ragionare per induzione sul rango per provare proprietà di elementi arbitrari dell'universo.

Il collegamento tra assioma di fondazione e principio di induzione non si ferma qui e ci porta direttamente al problema della categoricità. Come l'assioma di induzione – almeno al secondo ordine – garantisce la categoricità degli assiomi di Peano, così Fraenkel aveva cercato un analogo assioma nel caso della teoria degli insiemi,

l'assioma detto di *restrizione*, di cui quello di fondazione costituisce una sorta di indebolimento. Ruolo dell'assioma doveva essere quello di isolare tra tutti i possibili modelli una sorta di modello minimale, contenente solo quegli insiemi che, a partire dall'insieme vuoto, si potevano ottenere applicando le operazioni garantite dagli assiomi. Paradossalmente, è proprio con una tecnica analoga – modellata sulle catene di Dedekind – che Skolem aveva dimostrato il suo teorema sui modelli numerabili degli assiomi di Zermelo. Ciò che rendeva impossibile la formulazione di un assioma che svolgesse il ruolo richiesto da Fraenkel era proprio il carattere non assoluto di molte operazioni insiemistiche (tipo quella di passaggio alle parti) che Skolem aveva mostrato. Sarebbe stato von Neumann di lì a poco, proseguendo le indagini di Skolem, a cominciare un'analisi sistematica dei sottosistemi di modelli degli assiomi di Zermelo, mettendo in luce gli ostacoli enormi che si opponevano ad ogni tentativo di categorizzazione della teoria.

A prescindere dalle divergenze ideologiche, emergeva comunque, da questo intenso lavoro attorno agli assiomi di Zermelo, il profilo di un sistema assiomatico per la teoria degli insiemi che sarebbe divenuto paradigmatico – al contrario dei *Principia* – non solo per i filosofi ma anche per i matematici. È quel sistema $\mathfrak{ZF}$ (la notazione ormai in uso trascura inspiegabilmente il decisivo apporto di Skolem) che è formulato in un linguaggio del primo ordine, con la sola costante descrittiva $\in$ e che è oggi noto a tutti i ricercatori. Ne diamo qui di seguito una formulazione non molto diversa da quella che lo stesso Zermelo fornì nel 1930, accettando le critiche di Skolem. Si noti che Zermelo non includeva l'assioma di scelta, in quanto il suo carattere era diverso dagli altri assiomi non servendo a delimitare i modelli. Questo anche se Zermelo ovviamente accettava l'assioma, di cui proprio nello stesso articolo dedicato alla nozione di «definito» Fraenkel aveva dimostrato nel 1922 l'indipendenza dal sistema $\mathfrak{Z}$. Ma sui problemi dell'assioma di scelta avremo occasione di ritornare in modo più disteso nel capitolo VI.

$\mathfrak{ZF}$.1 (*Assioma di estensionalità*)

$$\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y$$

$\mathfrak{ZF}$.2 (*Assioma dell'insieme-coppia*)

$$\neg\, x = y \rightarrow \exists z\, \forall w\, (w \in z \leftrightarrow w = x \vee w = y)$$

$\mathfrak{ZF}$.3 (*Assioma dell'insieme-riunione*)

$$\exists y\, y \in x \rightarrow \exists z\, \forall w\, (w \in z \leftrightarrow \exists u\, (w \in u \wedge u \in x))$$

$\mathfrak{ZF}$.4 (*Assioma dell'insieme-potenza*)

$$\exists y\, \forall z\, (z \in y \leftrightarrow \forall w (w \in z \rightarrow w \in x)).$$

$\mathfrak{ZF}$.5 (*Schema d'assiomi di isolamento*)

Ogni espressione della forma:

$$\exists y\, \forall z\, (z \in y \leftrightarrow z \in x \wedge \mathscr{A}(z))$$

è un assioma purché $\mathscr{A}$ non contenga libera la variabile y.

$\mathfrak{ZF}$.6 (*Assioma di scelta*)

$$\forall y\, \forall z\, ((y \in x \wedge z \in x) \wedge \neg\, y = z) \rightarrow (\exists w (w \in y \wedge \neg\, \exists w w \in y \wedge w \in z))) \rightarrow \exists u\, \exists y\, (y \in x \rightarrow \exists w\, \forall v (v = w \leftrightarrow v \in u \wedge v \in y)).$$

$\mathfrak{ZF}$.7 (*Assioma dell'infinito*)

$$\exists z (\forall\, x (\neg\, \exists y\, y \in x \rightarrow x \in z) \wedge \forall x\, \forall y\, ((x \in z \wedge \forall w\, (w \in y \leftrightarrow w = x)) \rightarrow y \in z)).$$

$\mathfrak{ZF}$.8 (*Schema d'assiomi di rimpiazzamento*)

Ogni espressione della forma

$$\forall y\, \forall z\, \forall w\, (\mathscr{A}(y, z) \wedge \mathscr{A}(y, w) \rightarrow z = w) \rightarrow \\ \rightarrow \exists u\, \forall z\, (z \in u \leftrightarrow \exists y\, (y \in x \wedge \mathscr{A}(y, z)))$$

è un assioma purché $\mathscr{A}$ non contenga libere x e u.

$\mathfrak{ZF}$.9 (*Assioma di fondazione*)

$$\exists y\, y \in x \rightarrow \exists z\, (z \in x \wedge \forall w\, (w \in z \rightarrow \neg\, w \in x))$$

Vogliamo ora occuparci di un altro tipo di fondazione della teoria assiomatica degli insiemi che nasce in questo periodo ad opera di von Neumann, il quale negli anni venti dedicò alla questione almeno sei fondamentali lavori. In uno di questi, del 1925, *Eine Axiomatisierung der Mengenlehre* (*Una assiomatizzazione della teoria degli insiemi*) egli presentava un nuovo modo di assiomatizzare la teoria con notevoli caratteristiche di naturalezza in quanto evitava l'introduzione di alcune condizioni necessarie per il sistema di Zermelo ma tuttavia altamente non intuitive, prima fra tutte l'impossibilità, pena

il ricadere nelle contraddizioni, di parlare della classe totale e in generale di estensioni di proprietà che non siano insiemi.

Le classi hanno un ruolo centrale nella matematica ed in particolare nella teoria degli insiemi: si pensi alla classe degli insiemi fondati, dei cardinali, degli ordinali, ecc. Uno dei contributi più significativi dati da von Neumann in questi anni riguarda appunto il concetto di ordinale, che Cantor identificava con le classi di equivalenza di insiemi bene ordinati. Definiti in questi termini gli ordinali non potevano essere oggetto della teoria di Zermelo e comunque la loro totalità non poteva costituire una classe in quanto singolarmente essi sono classi. L'idea di von Neumann fu quella di sviluppare una teoria autonoma degli ordinali che risultasse uniforme per finiti e transfiniti, in cui la relazione d'ordine è la relazione di appartenenza $\in$. Un ordinale per von Neumann coincide con la famiglia dei suoi predecessori e formalmente si può definire come un insieme α ristretto al quale la relazione di appartenenza $\in$ risulta transitiva (ed α quindi è un insieme transitivo) e ben fondata. In presenza dell'assioma di fondazione quest'ultima clausola si può ridurre alla connessione, in base alla quale per ogni β e $\gamma \in \alpha$ si ha $\gamma \in \beta$, oppure $\beta \in \gamma$. In altre parole gli ordinali nel senso di von Neumann sono rappresentanti delle classi di equivalenza cantoriane, in cui la relazione d'ordine è data dall'appartenenza. Fatto importante, si può provare che la collezione *Ord* degli ordinali non può essere un insieme come pure la classe *Card* dei cardinali identificati con gli ordinali iniziali, quegli ordinali cioè che non sono equipotenti a nessun ordinale strettamente minore. In questo modo la teoria degli ordinali transfiniti ammette uno sviluppo autonomo indipendentemente dall'assioma di scelta e dal principio del buon ordinamento. Ma torniamo alla assiomatizzazione di von Neumann.

Von Neumann sviluppa il suo sistema avendo come sfondo la convinzione che le antinomie non traggono origine dall'assumere l'esistenza di un insieme in corrispondenza ad ogni proprietà – come appunto richiesto dall'assioma di comprensione – bensì dal fatto che tale insieme venga sostanzializzato, che gli si attribuiscano cioè quelle caratteristiche che possono essere riassunte semplicemente dicendo che l'insieme in questione può essere oggetto di predicazione, ossia può figurare come elemento di altre classi, molteplicità, insiemi. Ne viene allora che si potrà ammettere sen-

z'altro l'esistenza di una *molteplicità* in corrispondenza a ogni proprietà ma che occorrerà successivamente ricordare che non tutte queste molteplicità possono a loro volta essere *elementi* di altre molteplicità; in termini tecnici più precisi, occorrerà fare una netta e rigorosa distinzione fra quelli che oggi vengono chiamati *insiemi* e quelle che invece vengono chiamate *classi*: ogni insieme è una classe, ma non tutte le classi sono insiemi.[39] Nei termini che abbiamo sopra impiegato dovrebbe venir naturale esprimere la cosa dicendo che mentre tutti gli insiemi possono essere argomenti di predicazione ciò non vale per le classi; in termini più generali, esistono delle «cose» che non sono «argomenti».

Si comprende quindi come von Neumann conduca la sua assiomatizzazione in termini di funzioni piuttosto che di insiemi: la prima nozione è infatti abbastanza ampia da includere la seconda; più precisamente si tratta di due concetti del tutto equivalenti, dal momento che una funzione può essere riguardata come un insieme di coppie ordinate, e gli insiemi come funzioni particolari (funzioni caratteristiche) ossia come quelle funzioni che possono assumere due e solo due valori distinti. Afferma von Neumann: «La ragione per cui ci discostiamo dal modo usuale di procedere è che ogni assiomatizzazione della teoria degli insiemi usa la nozione di funzione (assioma di isolamento, assioma di rimpiazzamento...) e così è formalmente più semplice basare la nozione di insieme su quella di funzione che non viceversa.»

Fatte queste precisazioni, von Neumann prende le mosse dalla considerazione di due domini distinti, quello delle funzioni e quello degli argomenti e si chiede: quali funzioni sono nello stesso tempo argomenti? Osservando che qui «funzione» e «argomento» vanno intesi in senso puramente formale, ossia senza che venga loro attribuito uno specifico significato, si può pensare di considerare astrattamente «cose di tipo I» e «cose di tipo II» e chiedersi allora: quali cose sono di tipo I-II? (È chiaro che se ad esempio per «cose di tipo I» intendiamo «argomenti» e per «cose di tipo II» intendiamo «funzioni» porsi la domanda precedente significherà appunto chiedersi: quali funzioni sono anche argomenti? e quindi, per quanto già detto, quali classi sono anche insiemi?) Per rispon-

[39] Si noti che così facendo von Neumann riprende sostanzialmente la distinzione cantoriana fra molteplicità coerenti (insiemi) e molteplicità incoerenti (classi).

dere, con von Neumann, a questa domanda fondamentale, fissiamo arbitrariamente un argomento A (che avrà il ruolo dello 0 considerando le funzioni caratteristiche) e conveniamo che una funzione è anche un argomento quando, per così dire «*non lo è troppo spesso*» ossia quando non si comporta in modo tale da assumere «*per troppi argomenti*» un valore diverso da quello prefissato A. Poiché, prosegue von Neumann «un insieme sarà definito come una funzione che può assumere solo due valori, uno dei quali è A, questo è un adattamento ragionevole del punto di vista di Zermelo».

È chiaro così che il discorso di von Neumann non fa sostanzialmente che riprodurre, nella nuova terminologia, la constatazione che in un qualche senso (pena cioè il sorgere di antinomie) non possono esistere *insiemi* «troppo grandi» («funzioni troppo ampie» nella sua terminologia). L'idea centrale dell'assiomatizzazione di von Neumann riguarda il modo in cui viene precisato il concetto di «troppo grande» e si fonda essenzialmente sulla stessa idea che sta alla base dell'assioma di rimpiazzamento: immagini, rispetto a relazioni funzionali, di insiemi devono essere insiemi. Ciò significa che «grandi» sono quegli oggetti che si possono rappresentare mediante una relazione funzionale sull'oggetto grande per antonomasia: la classe totale. L'assioma centrale di von Neumann afferma quindi che un oggetto di tipo II è un oggetto di tipo I se e solo se non è rappresentabile mediante un oggetto di tipo II (una relazione funzionale) sulla classe totale. Il fatto estremamente interessante è che l'assioma di von Neumann non solo ha come conseguenza quello di rimpiazzamento – il che non sorprende – quanto anche l'assioma di scelta. Come avevamo detto sopra, infatti, gli ordinali costituiscono una classe *Ord* che corrisponde ad un oggetto di tipo II che non può essere di tipo I: ciò significa allora che esiste una corrispondenza tra l'universo e gli ordinali in forza dell'assioma. Disponiamo così di un buon ordinamento della classe totale e *a fortiori* ogni insieme sarà bene ordinabile.

Naturalmente, non è questo il solo assioma del sistema di von Neumann, che introduce ben cinque gruppi di assiomi: un primo gruppo, degli assiomi introduttivi, che riguarda essenzialmente il concetto di coppia ordinata e il concetto di applicazione, l'operazione cioè che ad ogni oggetto x di tipo II e ad ogni oggetto y di tipo I fa corrispondere un oggetto $[x, y]$ di tipo I (il valore della fun-

zione x applicata all'argomento y) ed include l'assioma di estensionalità; altri due gruppi detti rispettivamente degli assiomi aritmetici e degli assiomi logici di costruzione, che stabiliscono rispetto a quali operazioni sono chiusi tanto l'universo degli oggetti di tipo I quanto quello degli oggetti di tipo II. In forza di questi assiomi esistono ad esempio la funzione identica, le funzioni costanti, ecc., e l'universo è chiuso rispetto alle definizioni mediante specifiche forme di predicati. A questi seguono due altri gruppi, gli assiomi dell'infinito e gli assiomi delle cose di tipo I/II.

Gli assiomi dell'infinito corrispondono sostanzialmente a quello dell'infinito, della riunione e della potenza del sistema di Zermelo. Il loro nome scende dal fatto che secondo von Neumann essi sono necessari solo nel caso da una teoria generale degli insiemi e delle classi si voglia passare a una teoria del transfinito. L'altro gruppo contiene gli assiomi tipici del sistema: il primo ci garantisce sostanzialmente l'esistenza della classe di tutti gli insiemi, il secondo è quello che abbiamo sopra illustrato.

Rimane l'ultimo gruppo di assiomi, il cui ruolo è quello di restringere il più possibile la classe dei modelli, in vista di una qualche forma di categoricità; tra di essi – accanto agli assiomi che escludono l'esistenza di atomi e fissano l'interpretazione dell'oggetto A di cui si è detto sopra e di un correlato oggetto B come possibili valori di verità delle funzioni caratteristiche – si trova una formulazione forte dell'assioma di fondazione che riguarda classi e non solo insiemi.

Von Neumann non si limita a formulare il suo sistema e a mostrare come in esso si possa sviluppare la teoria cantoriana del transfinito ma, come abbiamo già ricordato, affronta sistematicamente uno studio dei possibili modelli del suo sistema, cercando di analizzare entro che limiti si possa sperare in qualche forma di categoricità. Von Neumann ha presente il risultato di Skolem sui modelli numerabili, che si estende anche al suo sistema, ed è scettico al riguardo. In particolare egli sviluppa un'analisi dettagliata dei sottosistemi dei modelli della teoria in collegamento con l'assioma di restrizione di Fraenkel; in questa analisi egli pone in luce le difficoltà di ottenere per via assiomatica solo modelli minimali e getta le basi di quella che sarà la tecnica dei modelli interni. La conclusione generale di von Neumann è non solo che distinzioni fondamentali riguardanti la cardinalità non hanno carattere asso-

luto (e questa è una ripresa del relativismo di Skolem) ma egli ribadisce che «abbiamo più di una ragione per avere riserve nei confronti della teoria degli insiemi» dal punto di vista fondazionale e che, allo stato dei fatti, non sembra ci sia via per riabilitarla.

Le ragioni della sfiducia di von Neumann poggiavano sulla radicale «non categorizzabilità» della teoria degli insiemi assiomatica in forza dell'idea che, se ha uno scopo assiomatizzare una teoria intuitiva come quella degli insiemi, questo deve essere la caratterizzazione univoca del modello intuitivo. Ben altre furono le conseguenze che nei primi anni trenta Zermelo avebbe tratto da questa situazione. Ben lungi dal ritenere la non categoricità un difetto della teoria, Zermelo da una parte poneva in luce come risultati forti come quello di Skolem andassero piuttosto visti come spie dei limiti del linguaggio adottato, quello del primo ordine, dall'altra ribadiva la intrinseca incompletabilità dell'universo degli insiemi, la sua potenziale estendibilità. Una volta costruita una totalità che può formare un modello degli assiomi dati, la logica interna della nozione di insieme ci porta a considerare questa stessa totalità come un insieme e a passare ad universi più grandi. In quest'ottica, in un fondamentale articolo del 1930 dal titolo *Über Grenzzahlen und Mengenbereiche* (*Numeri-confine e domini di insiemi*) Zermelo sfruttava la rappresentazione gerarchica degli universi data da von Neumann, indagando forme di categoricità relativa, individuando invarianti di isomorfismo nella cardinalità delle basi (corrispondente al livello zero dei modelli) e nell'altezza della scala degli ordinali contenuti nei modelli. Questo lo portava ad isolare il ruolo di quei cardinali inaccessibili che sarebbero poi divenuti, come vedremo, oggetto di studio approfondito da parte di numerosi logici. Nella contraddizione fra capacità espressiva dei linguaggi e principi di costruzione, Zermelo prendeva le parti dei secondi, mentre la ricerca logica nel suo complesso e l'indagine assiomatica sulla teoria degli insiemi avrebbero preso la strada indicata da Skolem.

È infatti utilizzando ancora i linguaggi del primo ordine che l'impostazione assiomatica di von Neumann verrà ripresa negli anni trenta da vari autori fra i quali Rafael Mitchel Robinson in *The Theory of classes, a modification of von Neumann's system* (*La teoria delle classi, una modificazione del sistema di von Neumann*, 1937) – che mantiene l'originale impostazione di von Neumann in termini di

funzioni – poi da Paul Bernays in una serie di lavori fra il 1937 e il 1958 a loro volta ripresi da Kurt Gödel in un suo fondamentale lavoro del 1940 che vedremo nel prossimo capitolo. Questi due ultimi autori ritornano tuttavia a un linguaggio più usuale (non in termini di funzioni cioè) mantenendo della teoria di von Neumann la distinzione classe/insieme. Sistemi di questo tipo vengono indicati nella letteratura come sistemi $\mathfrak{NBG}$ (von Neumann-Bernays-Gödel). Come esempio di sistema $\mathfrak{NBG}$ preferiamo qui riportare il sistema Σ impiegato da Gödel nel 1940, al quale dovremo riferirci in seguito.

Come Gödel stesso riconosce il suo sistema si collega essenzialmente a quello di Bernays ed è equivalente a una particolare versione del sistema originale di von Neumann. Da un punto di vista formale esso è formulato in un linguaggio del primo ordine con identità, con tre predicati extralogici, due monadici, Cx e Mx rispettivamente per «x è una classe» e «x è un insieme», uno diadico $x \in y$ per la relazione di appartenenza; per comodità di lettura assume due sorte di variabili: lettere *minuscole* per indicare insiemi e *maiuscole* per classi. Si hanno così cinque gruppi di assiomi che codificano le considerazioni che abbiamo visto sopra.

Il primo gruppo formula la distinzione fra insiemi e classi, fissando che insieme è una classe che è anche elemento; il secondo stabilisce invece criteri per la esistenza di classi e collettivamente equivale ad un principio di comprensione per formule $A(x)$ che siano *predicative*, non contengano cioè quantificazioni su variabili per classi. Il vantaggio della formulazione adottata da Gödel è che mentre il principio di comprensione formulato al primo ordine sarebbe uno schema, qui esso viene sostituito da otto condizioni di chiusura dell'universo delle classi, che ripercorrono la definizione induttiva di formula, in modo da garantire che per ogni $A(x)$ esiste la classe associata; mentre ad esempio B1 ci garantisce che esiste la classe delle coppie che costituiscono il grafo della relazione $\in$, B3 garantisce l'esistenza del complemento (che corrisponde alla negazione), B2 quella della intersezione (che corrisponde alla congiunzione) e B4 quella della proiezione parallela lungo un'asse (che corrisponde alla quantificazione esistenziale). I rimanenti assiomi del gruppo riguardano le permutazioni degli argomenti in una formula. Mentre i gruppi D ed E sono dati dall'assioma di

fondazione e da quello di scelta, il gruppo C riproduce gli assioni dell'infinito di von Neumann (infinito, riunione e potenza) e il principio di rimpiazzamento che è appunto un *assioma* (e non uno schema) in quanto è possibile quantificare su classi: esso afferma che per ogni classe A che rappresenti una relazione funzionale (in simboli, Un A) è un insieme l'immagine rispetto ad A di qualunque insieme x. Diversamente dal sistema $\mathfrak{ZF}$ il sistema di Gödel, che viene anche indicato con $\mathfrak{GB}$, risulta finitamente assiomatizzato, proprio per la presenza delle classi.

Gruppo A

A 1. Cx
A 2. $X \in Y \rightarrow MX$
A 3. $\forall x(x \in X \leftrightarrow x \in Y) \rightarrow X = Y$
A 4. $\forall x \, \forall y \, \exists z \, (u \in z \leftrightarrow (u = x \vee u = y))$

Gruppo B (assiomi delle classi)

B 1. $\exists A \, \forall x \, \forall y \, (<x, y> \in A \leftrightarrow x \in y)$
B 2. $\forall A \, \forall B \, \exists C \, \forall x \, (x \in C \leftrightarrow x \in A \wedge x \in B)$
B 3. $\forall A \, \exists B \, \forall x \, (x \in B \leftrightarrow \neg x \in A)$
B 4. $\forall A \, \exists B \, \forall x \, (x \in B \leftrightarrow \exists y \, (<y, x> \in A))$
B 5. $\forall A \, \exists B \, \forall x \, \forall y \, (<y, x> \in B \wedge x \in A)$
B 6. $\forall A \, \exists B \, \forall x \, \forall y \, (<x, y> \in B \leftrightarrow <y, x> \in A)$
B 7. $\forall A \, \exists B \, \forall x \, \forall y \, \forall z \, (<x, <y, z>> \in B \leftrightarrow <y, <z, x>> \in A)$
B 8. $\forall A \, \exists B \, \forall x \, \forall y \, \forall z \, (<x, <y, z>> \in B \leftrightarrow <x, <z, y>> \in A)$

Gruppo C (infinito, riunione, potenza, rimpiazzamento)

C 1. $\exists x \, (\exists y \, (y \in x) \wedge \forall z \, (z \in x \rightarrow \exists u \, (u \in x \wedge z \neq u \wedge \forall w \, (w \in z \rightarrow w \in u))))$
C 2. $\forall x \, \exists y \, \forall z \, \forall u \, (z \in u \wedge u \in x \rightarrow z \in y)$
C 3. $\forall x \, \exists y \, (\forall z \in u \rightarrow z \, (z \in x) \rightarrow u \in y))$
C 4. $\forall x \, \forall A \, (\text{Un} \, A \rightarrow \exists y \, \forall u \, (u \in y \leftrightarrow \exists v \, (v \in x \wedge <u, v> \in A)))$

Gruppo D (fondazione)

$\exists x \, (x \in A) \rightarrow \exists y \, (y \in A \wedge \forall u \neg \, (u \in y \wedge u \in A))$

Gruppo E (scelta)

$\exists A \, (\text{Un} \, A \wedge \forall x \, (\exists y \, (y \in x) \rightarrow \exists z \, (z \in x \wedge <z, x> \in A)))$.

Un problema si poneva naturalmente ed era quello dei rapporti fra i sistemi come $\mathfrak{ZF}$ e quelli dei tipo $\mathfrak{NBG}$. Come vedremo più avanti, negli anni fra il quaranta e il cinquanta, si otterranno risposte significative al riguardo e in particolare verrà dimostrato che $\mathfrak{NBG}$ è un'estensione conservativa di $\mathfrak{ZF}$ rispetto al linguaggio insiemistico. La situazione cambia se, come proposto da A. Morse alla fine degli anni quaranta, si rafforzano le ipotesi sulle classi e al posto del principio di comprensione e del gruppo B si assume uno schema che non si limita più alle formule predicative. Il sistema $\mathfrak{M}$ che ne risulta si dimostrerà particolarmente utile nella pratica matematica, prova ne sia che compare per la prima volta in appendice al volume di topologia di John L. Kelley.

3.3 *Logicismo: la «revisione» di Ramsey*

Come si ricorderà dal paragrafo 2.1, lo sviluppo che i *Principia* avevano dato del programma logicista lasciava aperti diversi problemi, di cui lo stesso Russell del resto era ben consapevole. Il primo riguardava la macchinosità di tutto l'apparato formale con la conseguente frattura fra la ricostruzione che i *Principia* offrivano della matematica e la matematica quale realmente veniva praticata. Questo in particolare riguardava la teoria cantoriana degli insiemi, e non stupisce che la maggior parte dei matematici – soprattutto tedeschi e francesi – si rifiutassero di vedere nei *Principia* un passo verso la fondazione di tale teoria come teoria matematica. Le simpatie andavano ai tentativi assiomatici come quello di Zermelo sino a contrapporre in modo esplicito – come faceva A. Schoenflies – nella crisi causata dalle antinomie, una dimensione puramente filosofica, di cui appunto si occupavano i *Principia*, ed una genuinamente matematica che di tutto l'apparato elaborato da Russell poteva fare a meno.

Senza giungere a questo atteggiamento estremo va osservato che diversi furono i tentativi di fornire un'analisi delle antinomie alternativa a quella di Russell, sia in direzione predicativa – come vedremo avanti accennando ai lavori di H. Weyl – sia in altre direzioni. In questo contesto acquista un particolare significato la ricezione che i *Principia* ebbero in Polonia, in un momento in cui, come abbiamo già ricordato, sia su un piano filosofico che su quello matematico, la logica e la teoria degli insiemi si andavano svi-

luppando in nuove direzioni. Si pensi, per far solo un esempio, che ai primi anni del secolo risale la nascita dei *Fundamenta Mathematicae*, una delle riviste più impegnate nella diffusione dell'approccio insiemistico all'Analisi, alla topologia, all'algebra, ecc. Abbiamo già accennato nel paragrafo 1 agli interventi di Chwistek relativamente al problema delle antinomie, ma il tentativo più radicale di ripensare nella sua interezza tale problematica nella Polonia di questi anni è dovuto sicuramente a Stanislaw Lesniewski che con Łukasiewicz fu uno dei fondatori della scuola di Leopoli.

A partire dal 1914 Lesniewski rifondò tutto intero l'edificio della logica e della teoria degli insiemi in una prospettiva che coniugava il logicismo russelliano con richiami alla fenomenologia di Husserl (non si dimentichi che Husserl era stato il maestro di Twardowski). Lesniewski articolava la sua ricostruzione in tre sistemi: la *prototetica*, che costituisce lo studio delle forme proposizionali ed è sostanzialmente una sorta di linguaggio proposizionale in cui è ammessa la quantificazione su variabili proposizionali, l'*ontologia*, il cui obiettivo è l'analisi della copula «è» e che ingloba in sé una logica paragonabile alla teoria dei tipi finiti, ed infine la *mereologia*, il cui obiettivo è lo studio della relazione parte/tutto. I sistemi di Lesniewski forniscono un'analisi decisamente alternativa a quella dei *Principia*, anche se in comune hanno l'obiettivo di una ricostruzione globale dell'universo logico.

La difficoltà più grossa che i *Principia* presentavano risiedeva però nel loro impianto filosofico. Russell e Whitehead sostenevano che la matematica intera faceva parte della logica e ciò, almeno nella prima edizione dei *Principia*, voleva dire che ogni enunciato matematico vero era traducibile in un linguaggio privo di costanti che non fossero logiche e dimostrabile a partire dagli assiomi dei *Principia*. In che senso però questo significa una riduzione alla logica? C'era una sottile ambiguità in questa concezione, che in parte, ma solo in parte, possiamo attribuire anche a Frege. Per Frege, come si ricorderà, le proposizioni logiche sono analitiche non nel senso che hanno una forma particolare, ma per il fatto di poter essere giustificabili senza ricorrere a considerazioni sintetiche: «il problema è», sosteneva Frege, «quello di trovare una prova della proposizione e di portarla avanti sino alle verità primitive. Se nel corso di questo processo incontriamo solo leggi logiche generali e definizioni, allora la verità è una verità analitica».

Se nel sistema originario di Frege però non era difficile dare un senso a questa locuzione, in quanto il sistema dei *Grundgesetze* sembrava non far altro che formulare le proprietà generali degli operatori logici coinvolti (anche il principio di comprensione non aveva una forma esistenziale, ma stabiliva le proprietà dell'operatore di astrazione) la cosa diventava più difficile per il sistema dei *Principia*: assioma moltiplicativo, dell'infinito e soprattutto assioma di riducibilità hanno una chiara forma esistenziale e soprattutto non hanno nessuna cogenza intrinseca sul piano dell'evidenza. La loro natura logica sta semmai nel fatto che riguardano un universo di oggetti logici, ed è questo che spiega l'atteggiamento sperimentale di Russell che poggia su una convinzione realista, per cui – come detto in una famosa pagina dell'*Introduzione alla filosofia matematica* – «la logica riguarda il mondo reale come la zoologia, anche se nei suoi aspetti più astratti e generali».

In questo senso la natura logica della matematica veniva a coincidere con la possibilità di ricostruzione all'interno dell'universo degli oggetti astratti. Russell stesso però era consapevole che non era sufficiente qualificare come logico tutto ciò che è vero ed universale e che doveva esserci qualcosa in più per isolare ciò che è logico da ciò che non ha controesempi. È in questo senso che nell'*Introduzione alla filosofia matematica* Russell comincia a parlare di carattere tautologico delle leggi logiche e attribuisce a Wittgenstein il merito di avere introdotto questa precisazione. «Tutte le proposizioni della logica» egli scrive «hanno una caratteristica che si esprimeva di solito dicendo che sono analitiche, o che le loro contraddittorie sono autocontraddittorie. Questo modo di formulare la cosa però non è soddisfacente. La legge di contraddizione è solo una fra le tante proposizioni logiche; non ha preminenza speciale; e la prova che la contraddittoria di qualche proposizione è autocontraddittoria richiede spesso altri principi di deduzione oltre alla legge di contraddizione. Ciononostante la caratteristica delle proposizioni logiche che noi cerchiamo fu intuita e si cercò di definirla da parte di coloro che dicevano che consisteva nella deducibilità dalla legge di contraddizione. Questa caratteristica ... per il momento la possiamo chiamare tautologicità».

Russell non riusciva però a spiegare in che cosa consistesse la tautologicità e qui stava forse la difficoltà più profonda che la sua versione del logicismo incontrava: è bizzarro voler fondare la ma-

tematica sulla logica e non essere in grado di dire che cosa è la logica. Fu Wittgenstein che nel *Tractatus logico-philosophicus* (*Trattato logico-filosofico*, 1922) e in altri scritti più o meno contemporanei cercò nel modo più deciso di superare questa difficoltà e di sottolinearne la rilevanza. Al di là di analisi dettagliate che non possiamo svolgere in questo contesto, si può dire che l'idea di fondo di Wittgenstein stava nel generalizzare a tutta intera la logica la caratteristica che il metodo delle tavole di verità di Peirce-Schröder aveva posto in luce nel caso proposizionale: tautologia è tutto ciò che è vero in qualunque stato possibile di cose (per qualunque assegnazione di valori di verità alle formule atomiche) e che quindi esaurisce tutte le possibilità, così che la sua negazione diviene automaticamente autocontraddittoria. Per Wittgenstein questo voleva dire che le tautologie non hanno significato; basandosi sulla sua distinzione fra «mostrare» e «dire»: un linguaggio logico perfetto avrebbe dovuto *mostrare* automaticamente la tautologicità delle leggi logiche mediante una semplice ispezione della loro forma ed in questo senso per molto tempo per Wittgenstein la possibilità di definire il concetto di tautologia come legge logica fu legato a quello di trovare un metodo di decisione.

In questa prospettiva tutta intera la logica dei *Principia* doveva essere ridotta a forme proposizionali, sfruttando eventualmente sviluppi infinitari. In altre parole il linguaggio doveva essere traducibile in termini di combinazioni di connettivi (eventualmente infinitari) a partire da formule atomiche. Non ha senso qui raccontare in dettaglio le difficoltà che Wittgenstein incontrò nell'affrontare questo progetto e ci preme piuttosto sottolineare la portata generale delle questioni così sollevate: da una parte emergeva in primo piano il problema di elaborare una semantica in base alla quale definire il concetto di validità logica, che fosse più adeguata di quella in termini di «descrizioni di stato» data da Wittgenstein che sostanzialmente assume l'unicità dell'universo del discorso; dall'altra si poneva la questione di come, in questa versione più puntuale di logica, avesse senso il programma logicista. È su questo sfondo che emerge in primo piano la figura di un matematico, economista, logico e filosofo di marca dichiaratamente logicista, i cui possibili ulteriori contributi furono purtroppo tragicamente bloccati da una morte prematura. Alludiamo a Frank Plumpton Ramsey che, come dicevamo, muove una critica

«interna» al logicismo e, come dichiara esplicitamente nel saggio *The foundations of mathematics* (*I fondamenti della matematica*) del 1925, si propone di «fornire un'esposizione soddisfacente dei fondamenti della matematica secondo il metodo generale di Frege, Whitehead e Russell». Come Ramsey afferma, «seguendo questi autori io sostengo che la matematica è parte della logica, ed appartengo così a quella che si potrebbe chiamare la scuola logicista in contrapposizione alla scuola formalista e a quella intuizionista».

Oltre che per il suo tentativo di riproporre il logicismo all'attenzione degli studiosi, Ramsey è interessante anche per la valutazione che dà della situazione delle ricerche sui fondamenti di questo decennio. Quando ad esempio riconosce che il logicismo è in declino, è esplicito: «Ho quindi preso i *Principia mathematica* come base per la discussione e per una revisione; e credo di aver scoperto come, usando l'opera di Ludwig Wittgenstein, li si possa liberare dalle gravi obiezioni che li hanno fatti respingere dalla maggioranza delle autorità tedesche in materia, che hanno abbandonato l'approccio russelliano ai problemi della logica.» La matematica pura, sostiene Ramsey, può essere riguardata dal punto di vista delle sue idee o concetti o dal punto di vista delle sue proposizioni. Distinzione essenziale dal momento che la grande maggioranza degli studiosi «hanno concentrato l'attenzione sulla spiegazione dell'una o dell'altra di queste categorie supponendo erroneamente che alla soddisfacente spiegazione dell'una sarebbe immediatamente seguita quella dell'altra». Così i formalisti (Hilbert e la sua scuola) hanno concentrato la loro attenzione sulle proposizioni della matematica col risultato di trarne una teoria «disperatamente inadeguata»: i concetti della matematica infatti possono essere chiariti proprio dal fatto che essi «si presentano al di fuori della matematica, nelle proposizioni della vita quotidiana»; da parte loro, gli intuizionisti sostengono una teoria che comporta «dichiaratamente la rinuncia a molti dei più fruttuosi metodi dell'Analisi moderna, per l'unica ragione, mi sembra, che tali metodi mancano di conformarsi ai loro pregiudizi privati».[40]

[40] Per inciso, questo giudizio pesantemente ironico sugli intuizionisti, mostra bene, da una parte, l'incomprensione che in quel periodo regnava verso di loro, ma mette anche in evidenza, dall'altra, come fosse ormai necessario rendere «confrontabili» metodi e risultati di questa scuola con quelli delle altre.

Ma veniamo all'intervento di Ramsey sulla sistemazione dei *Principia*. I logicisti si sono concentrati sull'analisi dei concetti della matematica e li hanno ricondotti alla logica: ne hanno dedotto che le proposizioni matematiche sono quelle proposizioni vere in cui compaiono solo concetti matematici o logici, giusta la definizione russelliana dei *Principi della matematica*. Ma la questione non è ancora risolta, a parere di Ramsey, perché Russell è «*ancora ben lontano da un'adeguata* concezione della natura della logica simbolica a cui era stata ridotta la matematica». Ora, contrariamente a Russell che riteneva caratteristica delle proposizioni matematiche semplicemente la «completa generalità» senza indagare ulteriormente come i contenuti matematici delle proposizioni vengono giustificati sul piano della forma logica, Ramsey ritiene viceversa che il contenuto delle proposizioni matematiche «deve essere completamente generale e la loro forma deve essere tautologica». È chiara l'influenza del pensiero di Wittgenstein e la contrapposizione non potrebbe essere più netta: «i formalisti trascuravano completamente il contenuto e rendevano la matematica senza significato, i logicisti trascuravano la forma e facevano consistere la matematica in qualunque generalizzazione vera; solo tenendo conto di entrambi i punti di vista e considerando la matematica come composta di generalizzazioni tautologiche noi possiamo ottenere una teoria adeguata».

Ramsey ritiene, rafforzando se possibile la tesi di Wittgenstein, che la matematica consista interamente di *tautologie*. Se per matematica intendiamo la trascrizione che di essa si fa nei *Principia*, allora si tratterà di far vedere che le stesse proposizioni fondamentali dei *Principia* sono tautologie, dal momento che le regole di deduzione ammesse conducono da tautologie a tautologie. Eliminiamo subito il caso degli assiomi logici (e delle regole logiche). Per Ramsey le tautologie possono avere infiniti argomenti (viste come funzioni di verità) sicché i quantificatori $\forall$, $\exists$ si riducono a congiunzioni infinite ($\forall x \mathscr{A}$ diviene $\mathscr{A}(a_2) \wedge \mathscr{A}(a_2) \wedge \ldots$ per ogni individuo a_i del dominio) e a disgiunzioni infinite ($\exists x \mathscr{A}$ diviene $\mathscr{A}(a_1) \vee \mathscr{A}(a_2) \vee \ldots$ come sopra). Significativamente compaiono anche in questo caso, come in quello di Zermelo, formule infinite nel tentativo di dipanare i problemi connessi con la teoria delle classi. La grande difficoltà è data, come è ovvio, da una parte dall'assioma di riducibilità, che risulta chiaramente essere

una proposizione contingente, ossia né una tautologia né una contraddizione; dall'altra dalle assunzioni esistenziali che prospettano difficoltà che a parere di Ramsey non sono intrinseche e dipendono piuttosto dal modo di trattare alcuni concetti fondamentali nello sviluppo dei *Principia.* Ramsey affronta quindi questo problema avendo come mira quella di ridurre un calcolo estensionale come quello di Russell e Whitehead a un calcolo di funzioni di verità. Quali sono allora i difetti che egli ravvisa nei *Principia* e che vuole evitare grazie alla teoria delle tautologie? Essenzialmente i tre seguenti.

1) La teoria dei *Principia* non consente di considerare tutte le classi infinite poiché pretende che ogni classe sia definita da una funzione proposizionale, sicché non possiamo occuparci di classi (o insiemi) infinite, se ne esistono, che non siano definibili o definite da funzioni proposizionali. D'altra parte, osserva Ramsey, anche se non possiamo definirle, è chiaro che a queste classi «indefinibili» si fa necessariamente riferimento in locuzioni quali «per ogni classe» o «esiste una classe tale che» ossia in espressioni quantificate rispetto a simboli di classe. Potremmo certo convenire di limitare il nostro universo del discorso in modo da escludere le classi «indefinibili»; ma così facendo altereremmo in modo essenziale il senso delle locuzioni precedenti, adottando un atteggiamento predicativista in antitesi al realismo russelliano e alla pratica matematica. D'altra parte è chiaro anche che non sarebbe corretto, pena la distruzione della «apriorità e necessità che sono l'essenza della logica», limitarsi a considerare solo classi definibili. Il primo grave difetto dei *Principia* è proprio quello di non aver considerato la possibilità che esistano classi indefinibili, contravvenendo così, tra l'altro, all'atteggiamento estensionale della matematica moderna. L'errore consiste «nel dare una definizione di classe che si applica solo alle classi definibili, di modo che tutte le proposizioni matematiche su alcune o tutte le classi vengono interpretate in modo scorretto». A questo primo errore è direttamente legato il problema relativo all'assioma moltiplicativo; esso infatti «se rettamente interpretato è una tautologia, ma frainteso al modo dei *Principia mathematica* diventa una proposizione di significato empirico che non vi è ragione di supporre vera».

2) Il secondo difetto dei *Principia* è una visione troppo indiscriminata delle antinomie. Abbiamo già descritto come Russell pen-

sasse di superarle tramite la teoria ramificata dei tipi e abbiamo già notato come in effetti quest'ultima consista di due parti che possiamo chiamare la tipizzazione vera e propria e la ramificazione (ordini). Ora, osserva Ramsey, «queste due parti vennero unificate perché erano entrambe dedotte, in modo alquanto approssimativo, dal "principio del circolo vizioso"». La soluzione che Ramsey propone consiste nel considerare le antinomie distinte in due gruppi. Il gruppo A comprende l'antinomia di Russell, di Burali-Forti, di Cantor ecc. fino alla n. 5, mentre il gruppo B contiene l'antinomia del mentitore, l'antinomia di Richard, ecc., secondo l'elencazione che delle antinomie abbiamo fatto nel paragrafo 2.1. La distinzione tra i due gruppi avviene sulla base dei seguenti criteri. Le antinomie del gruppo A (che oggi vengono dette, dopo Ramsey, antinomie logiche) sono tali che «se non si prendessero provvedimenti contro di esse si presenterebbero negli stessi sistemi logici o matematici». In esse, si noti, intervengono solo concetti logici o matematici e il loro presentarsi significa appunto che ci deve essere «qualcosa di sbagliato» nella logica o nella matematica che noi adottiamo. Le antinomie del gruppo B, viceversa, «non sono puramente logiche e non possono venir enunciate in soli termini logici; poiché tutte contengono qualche riferimento al pensiero, al linguaggio o al simbolismo che non sono termini formali, ma empirici».[41]

È implicita nel discorso di Ramsey quella distinzione linguaggio/metalinguaggio che oggi è generalmente accettata come elemento fondamentale di chiarificazione nella costruzione dei sistemi formali. In questa prospettiva, la ramificazione dei tipi, dalla quale come sappiamo dipende l'assunzione dell'assioma di riducibilità russelliano, potrà essere evitata *come parte del sistema*, ossia apparterrà alla sfera *metalinguistica* del sistema stesso, nel quale verrà invece mantenuta la tipizzazione semplice.

3) Il terzo grave difetto che Ramsey – sulla scorta di Wittgenstein – riscontra nei *Principia* è connesso con la definizione di identità. Russell – come abbiamo notato – faceva dipendere essenzialmente questa definizione dall'assioma di riducibilità, in quanto assumeva per definizione che due «cose» sono identiche se e solo

[41] È a questo punto che Ramsey accredita a Peano l'aver riconosciuto tale distinzione. Ne abbiamo parlato, come il lettore ricorderà, alla fine del paragrafo 2.4.

se hanno in comune le proprietà *predicative* (dal che poi l'assioma garantiva l'estendibilità a tutte le proprietà). Ora Ramsey ritiene che questa dipendenza non esista e che il vero difetto di questa definizione sia analogo a quello già riscontrato per le classi al punto 1): si tratta cioè di una interpretazione scorretta dell'identità «in quanto non si definisce il significato per cui il simbolo dell'identità viene effettivamente usato». Come il primo difetto aveva un'immediata influenza sull'assioma moltiplicativo, questo relativo all'identità riguarda direttamente l'assioma dell'infinito, in quanto la cardinalità di un insieme dipende dal numero di oggetti *distinti* che contiene e quindi dalla verità o meno di enunciati del tipo $a \neq b$.

Vediamo ora brevemente come Ramsey ritiene di poter superare queste difficoltà presentate dai *Principia* per poter ancora sostenere la possibilità di realizzare il programma logicista. Il punto centrale della sua revisione consiste in una nuova definizione di *funzione predicativa* (che non va confusa con quella di Russell, che in questo contesto Ramsey chiama *elementare*). Ramsey parte dalla considerazione che l'essere una proposizione elementare o no (ossia predicativa o no nel senso di Russell) non è una caratteristica *reale* della proposizione stessa, bensì del suo *modo di espressione*, vale a dire, una stessa proposizione può essere espressa sia come proposizione elementare che come proposizione non elementare. Il punto è separare l'aspetto linguistico da quello reale: da una parte c'è il linguaggio, che «non è fissato oggettivamente ma dipende dai nostri metodi di costruire simboli», dall'altra le proposizioni e le funzioni su proposizioni che hanno una realtà oggettiva. È su di esse come sulle classi contrapposte alle funzioni proposizionali che per Ramsey bisogna porre l'enfasi. Per definire questo campo di simboli si può seguire il metodo di Russell, che Ramsey chiama *soggettivo*, ossia si può definire il campo delle funzioni che possono essere costruite in un certo modo (ad esempio con il semplice uso di un particolare connettivo) oppure si può usare quello che Ramsey chiama il metodo *oggettivo* e che fa suo; questo secondo modo consiste grosso modo nel trattare le funzioni di funzioni esattamente come si trattano le funzioni di individui; in altri termini, Ramsey propone di determinare i simboli che possono essere posti ad argomento di una funzione di funzione «non secondo il metodo della loro costruzione, ma secondo il loro significato». Con queste

premesse, Ramsey introduce il concetto di funzione predicativa intendendo con ciò «una funzione che è una qualsiasi *funzione di verità* di argomenti che, finiti o infiniti in numero, sono tutti o funzioni atomiche di individui o proposizioni» [corsivo nostro]. Si noti che questa definizione dipende in modo essenziale dal concetto di funzione di verità di un numero infinito di argomenti e quindi comprende in modo proprio le funzioni dei *Principia* che come si ricorderà potevano riguardarne solo un numero finito. È proprio qui che c'è il salto fra il metodo «costruttivo» dei *Principia* e il metodo «dei significati» di Ramsey.

Ramsey mostra che questa sua nozione di funzione predicativa da un lato permette di tener conto di ogni funzione (anche cioè di quelle che non possiamo effettivamente esprimere data la finitezza dei nostri mezzi linguistici) e d'altro lato non porta a contraddizioni. Ciò premesso Ramsey giunge a una gerarchia per tipi e ordini delle funzioni, dove, come al solito, i tipi riguardano gli argomenti mentre, indipendentemente dal tipo, l'ordine riguarda i valori, ma sottolinea «l'essenziale distinzione fra ordine e tipo: il tipo di una funzione è una sua caratteristica reale che dipende dagli argomenti che la funzione può assumere; ma l'ordine di una proposizione o di una funzione non è una caratteristica reale, bensì quella che Peano chiama una pseudofunzione». A questo punto si dimostra in modo molto simile a quello dei *Principia* che vengono evitate anche le contraddizioni del gruppo B; ma l'apparente analogia non deve ingannare perché, ribadisce Ramsey, «per me le proposizioni in se stesse non hanno ordine; sono solo differenti funzioni di verità di proposizioni atomiche – una totalità definita che dipende solo da quante proposizioni atomiche ci sono. Gli ordini e le totalità illegittime intervengono solo con i simboli che usiamo per simbolizzare i fatti in modi variamente complicati». Col che viene messo in evidenza in che senso il gruppo B di contraddizioni dipenda essenzialmente dalla componente linguistica.

Abbiamo sopra accennato all'esistenza di precisi rapporti, secondo Ramsey, di queste tre lacune dei *Principia* con quelli che sono gli assiomi «critici» della sistemazione di Russell e Whitehead e cioè con l'assioma di riducibilità, l'assioma moltiplicativo e l'assioma dell'infinito rispettivamente. Nella nuova visione di Ramsey la prima difficoltà viene eliminata nel modo più radicale perché la sua introduzione delle funzioni predicative e in estensione

rende semplicemente superfluo l'assioma di riducibilità, eliminando come fa dal contesto interno del sistema la ramificazione per ordini che viene a essere una sovrastruttura *esterna* alla costruzione stessa. Concependo le classi nel suo senso Ramsey mostra come l'assioma moltiplicativo diventi una tautologia, mentre l'accettazione dell'assioma dell'infinito viene a dipendere dal fatto che, nella interpretazione di Ramsey, esso viene ora a esprimere semplicemente che esiste un numero *infinito di individui* e non, come in Russell e Whitehead, che esiste un numero infinito di individui *distinguibili*. Grazie all'interpretazione mutuata da Wittgenstein ne risulta che l'assioma stesso è una tautologia o una contraddizione. In altri termini noi «non possiamo dire nulla» per quanto riguarda il numero degli individui che esistono perché ogniqualvolta tentiamo di farlo scriviamo o una tautologia o una contraddizione: possiamo fare solo asserzioni numeriche su universi limitati, non su tutto l'universo della logica.

Dopo questa sommaria presentazione del tentativo effettuato da Ramsey per il recupero della posizione logicista, è opportuno concludere con qualche osservazione, relativa allo stato generale della ricerca sui fondamenti di questo periodo, e quindi al rapporto delle varie scuole fra loro. Almeno in un primo momento Ramsey ritenne ovviamente corretta la tesi logicista, ove su di essa si intervenga con le vedute di Wittgenstein per quanto riguarda quella che abbiamo chiamato la tesi dell'estensionalità. Da una parte infatti egli non riusciva a capacitarsi del rifiuto intuizionista del terzo escluso (né nella forma radicale e diretta di Brouwer, né in quella più mediata, ma a suo parere altrettanto mutilante per la matematica, che ne dà Weyl). Né ammetteva l'eliminazione delle proposizioni generali ed esistenziali dal dominio delle proposizioni direttamente interpretabili operata da Weyl e alla quale si associa Hilbert. La concezione intuizionista che esse non siano affatto proposizioni e che quindi abbisognino di un substrato costruttivo, o quella formalista secondo la quale in queste proposizioni vanno individuati gli «elementi ideali» della nostra espressione linguistica della matematica non lo soddisfano; ad esse oppone la tesi di Russell (corroborata però dalla visione di Wittgenstein) secondo la quale *tutte* le proposizioni esprimono accordo o disaccordo con possibilità di verità di proposizioni atomiche, o in altri termini, sono funzioni di verità di proposizioni atomiche. Da questo punto di

vista, dal momento che una proposizione universale o esistenziale esprime solo il fatto che una almeno, o tutte le proposizioni considerate siano vere, non ha alcun valore o alcun peso la circostanza che questo insieme di proposizioni sia finito o infinito.

Tuttavia, e in particolare discutendo lo *status* dell'assioma dell'infinito, Ramsey concludeva con una nota di pessimismo. A suo parere il problema era posto in questi termini: una giustificazione logica della matematica pura è assai problematica perché «... Brouwer e Weyl dicono che ciò non è possibile e Hilbert propone di giustificarla come un gioco condotto sulla carta, con segni privi di significato»; d'altra parte riteneva che il suo stesso tentativo di ricostruire la teoria di Whitehead e Russell «... superi molte difficoltà, ma... è impossibile considerarlo come del tutto soddisfacente». E questa accennata insoddisfazione si trasforma in un graduale allontanamento dalla posizione logicista. Nel 1929 con l'articolo *Mathematics, foundations of* (*Matematica, fondamenti della*) scritto per la 14-esima edizione dell'*Enciclopedia britannica* si può notare un suo ulteriore avvicinamento al finitismo hilbertiano. Va del resto notato che già nell'articolo *On a problem of formal logic* (*Su un problema di logica formale*) del 1928 egli aveva preso in considerazione un tipico problema dell'indirizzo formalistico, il *problema della decisione* del quale aveva risolto un caso particolare facendo ricorso a un risultato assai profondo, oggi noto appunto come teorema di Ramsey, che interviene in numerosi e importanti campi della ricerca contemporanea.[42]

[42] Una formulazione del teorema di Ramsey può essere data come segue. Dato un insieme A, indichiamo con $P_2(A)$ l'insieme di tutti i sottoinsiemi di A che contengono esattamente due elementi (coppie *non* ordinate). Allora il teorema afferma che data una scomposizione di $P_2(A)$ in due insiemi M e N, $P_2(A) = M \cup N$, se A è infinito (il caso finito non offre alcuna difficoltà) è sempre possibile trovare un sottoinsieme infinito A_0 di A tale che $P_2(A_0) \subseteq M$ o $P_2(A_0) \subseteq N$. Ramsey in effetti ne dimostra una versione generalizzata nel senso che assume un numero finito qualsiasi di insiemi «scomponenti» A e i sottoinsiemi di A da lui considerati possono avere qualunque cardinalità finita purché prefissata (considera cioè P_n (a) per n finito prefissato qualunque). Come si nota il teorema è una generalizzazione del famoso teorema dei cassetti di Dirichlet, secondo il quale se un insieme infinito viene scomposto in un numero finito di insiemi A_1, A_2, ... A_n, ossia $A = A_1 \cup A_2 \cup \ldots \cup A_n$, allora almeno uno degli A_i è infinito. Malgrado l'apparente semplicità, il teorema di Ramsey richiede per la sua dimostrazione il ricorso all'assioma di scelta. Oltre all'applicazione indicata da Ramsey stesso, il teorema ne ha di recente ottenute numerose altre in particolare nell'ambito della teoria

Siamo così giunti in modo naturale a parlare di quell'indirizzo che negli anni venti sembra senza dubbio primeggiare, e avere le migliori credenziali per risolvere definitivamente il problema dei fondamenti della matematica. Alludiamo alla corrente formalista (anche se formalisti contemporanei, ad esempio Haskell B. Curry, propongono di chiamarla, più precisamente, hilbertismo) e il primo nome che qui si presenta naturale è ovviamente quello di David Hilbert, di cui abbiamo già considerato le prime prese di posizione. In particolare si ricorderà la polemica con Poincaré, originata dalla memoria hilbertiana del 1904, a proposito del principio di induzione. Dopo un lungo periodo di silenzio su questi temi, Hilbert riprese la questione dei fondamenti a partire dal 1917 con *Axiomatische Denken* (*Il pensiero assiomatico*), una conferenza tenuta a Zurigo, e in tutta una serie di conferenze e scritti venne precisando sempre più durante questo periodo quelli che sono i tratti fondamentali e caratteristici della sua scuola. Ricordiamo fra gli altri *Neubegründung der Mathematik (erste Mitteilung)* (*Nuova fondazione della matematica; prima comunicazione*, da una conferenza a Copenaghen-Amburgo, 1922), *Die logischen Grundlagen der Mathematik* (*I fondamenti logici della matematica*, 1923, da una conferenza a Lipsia del 1922), *Über das Unendliche* (*Sull'infinito*, 1926, da una conferenza a Münster del 1925), *Die Grundlagen der Mathematik* (*I fondamenti della matematica*, 1927, da una conferenza ad Amburgo), *Probleme der Grundlegung der Mathematik* (*Problemi della fondazione della matematica*, 1928, da una conferenza a Bologna) e dello stesso anno, in collaborazione con Ackermann, i *Grundzüge*, di cui si è parlato, ove veniva codificata quella classificazione dei linguaggi e dei calcoli logici che è ancor oggi corrente.

Particolarmente importanti la conferenza sull'infinito, nella quale si ha forse la più completa enunciazione del programma hilbertiano, come pure il susseguente articolo del 1927, a sfondo più polemico nei riguardi degli intuizionisti. Oltre alla versione organica della logica dei *Grundzüge*, Hilbert aveva già in preparazione

dei modelli, dell'Analisi funzionale e della combinatoria infinita, tanto che – in questi ultimi anni – si è venuto configurando tutto un campo di ricerca nota come Teoria di Ramsey.

verso la fine del decennio quella che sarà l'esposizione cardinale della *Beweistheorie*, che si concretizza nei due volumi delle *Grundlagen der Mathematik* (*Fondamenti della matematica*), scritti in collaborazione con Bernays e che vedranno la luce nel decennio successivo, nel 1934 e 1939 rispettivamente. Oltre ai diretti collaboratori di Hilbert già ricordati, è obbligo menzionare in questo contesto autori come Herbrand e Skolem – che hanno nella loro produzione o nel loro atteggiamento di fondo vari punti di contatto con la posizione hilbertiana.

Una soddisfacente esplicazione del problema dei fondamenti della matematica deve necessariamente passare, a parere di Hilbert, attraverso l'analisi e la chiarificazione di quello che è il concetto cardine, la nozione critica in tutto lo sviluppo della matematica: l'infinito. L'Analisi lo aveva solo apparentemente eliminato dalla sua problematica specifica, dopo che da essa erano state bandite le vaghe idee connesse con la nozione di infinitesimo; tuttavia l'infinito rientra ancora nella formulazione delle proposizioni matematiche (in particolare della stessa Analisi) ogniqualvolta si hanno espressioni che fanno riferimento a *tutti* i numeri reali o si afferma che *esistono* numeri reali che godono di date proprietà. Se chiamiamo allora problema dell'infinito quello che appunto consiste nella chiarificazione ed eliminazione di questo concetto dalla matematica, vediamo che tale problema, malgrado gli sforzi degli analisti dell'Ottocento, permaneva ancora aperto alla critica moderna. Fine della ricerca di Hilbert[43] è quello di far vedere come sia altrettanto una pura *façon de parler* l'infinito nel senso delle totalità infinite. Anche la via di soluzione del problema è indicata dal processo di rigorizzazione dell'Analisi: come allora si era riusciti a ridurre a operazioni col finito le operazioni con gli infinitesimi, così oggi «*i modi di inferenza che impiegano l'infinito devono in generale essere sostituiti con processi finiti che danno precisamente lo stesso risultato*» [corsivo nostro]. È questo il proposito della teoria di Hilbert, che si pone così, a suo parere, come naturale continuazione dell'opera di Weierstrass. È, come ognuno vede, l'enunciazione del programma o comunque dell'esigenza *finitista* che caratterizza appunto il formalismo hilbertiano. Il massimo tribunale posto a giudicare la

[43] E non solo, in particolare, del discorso *Sull'infinito*, ma della concezione globale di Hilbert.

correttezza (non, si badi bene, la fecondità) dei nostri comportamenti deduttivi è l'assenza di contraddizione: pena lo scadere in una vuota metafisica, è necessario concedere, secondo Hilbert, che una volta che una nozione sia stata introdotta e non porti a contraddizione tale nozione è acquisita per la matematica e si può fare *come se* essa avesse una realtà oggettiva alla Vahinger. E ciò comporta che «una chiarificazione definitiva della *natura dell'infinito* è divenuta necessaria, non semplicemente per l'interesse particolare delle singole scienze, ma piuttosto per *l'onore dell'intelletto umano stesso*».

Come dicevamo, secondo Hilbert, malgrado scienze naturali quali la fisica, o matematiche come la geometria, portino alla conclusione probabile che la realtà è finita, tuttavia ciò non toglie che «l'infinito abbia un posto ben giustificato nel *nostro pensiero*» e quindi assuma l'aspetto di un concetto indispensabile. Esempi illuminanti in questo senso sono dati dall'algebra, dalla geometria proiettiva e in particolare dall'Analisi che, a ben guardare, è null'altro che «una sinfonia dell'infinito». Ma la teoria che ci permette di penetrare più a fondo il significato della nozione di infinito è senza dubbio la teoria cantoriana degli insiemi, e in particolare la teoria cantoriana dei numeri transfiniti. Hilbert accetta ovviamente la distinzione tra infinito attuale e infinito potenziale e, con Cantor, ritiene che solo col primo abbiamo a che fare col «vero» infinito. Ma il presentarsi delle antinomie impone l'elaborazione di una nuova teoria, nella fattispecie la *Beweistheorie*, che superi le difficoltà.

Le motivazioni e i criteri sulla base dei quali Hilbert costruisce il proprio programma sono duplici:

1) è necessario indagare con cura i modi di costituire concetti e i metodi di inferenza fruttuosi nella pratica matematica;

2) occorre rendere le inferenze quanto più affidabili («sicure») è possibile ed indagare su che basi ciò può avvenire.

Ora, tutto ciò si può ottenere solo a patto di riuscire veramente a chiarire la nozione di infinito. Hilbert accetta quindi un primo principio direttivo, secondo il quale, in accordo con Kant, «la matematica dispone di un contenuto certamente indipendente da ogni logica e non si può quindi assicurarle una fondazione con i soli mezzi della logica». Abbiamo già visto che con la teoria della dimostrazione si trasforma ogni proposizione matematica in una

formula che può essere esibita concretamente; e ciò che in questo procedimento va assunto come dato «sono certi oggetti extralogici che sono intuitivamente presenti come esperienza immediata antecedente a ogni pensiero»; sicché nella teoria della dimostrazione «gli assiomi e le proposizioni dimostrabili... sono copie dei pensieri che costituiscono la matematica ordinaria come sviluppata sino a oggi». Ricordando le parole che lo stesso Hilbert aveva usato nella conferenza di Parigi nel 1900: «I segni aritmetici sono figure scritte e le figure geometriche sono formule disegnate», in entrambi i casi oggetti concreti. Orbene, se l'inferenza logica deve dare affidamento, deve essere possibile «osservare questi oggetti completamente in tutte le loro parti, e il fatto che essi occorrano, e che essi differiscano l'uno dall'altro e che si susseguano, o siano concatenati, è immediatamente dato in modo intuitivo, assieme agli oggetti, come qualcosa che né può essere ridotta a qualcos'altro, né richiede tale riduzione. *Questa è la posizione filosofica fondamentale, il requisito per la matematica e, in generale, per ogni pensiero, comunicazione, comprensione scientifici»* [corsivo nostro].

Certo la teoria dei numeri ad esempio può essere sviluppata solo sulla base di considerazioni contenutistiche riguardanti combinazioni finite di oggetti concreti, ma è altrettanto certo che queste non esauriscono la matematica. Tuttavia si può pensare di ridurre la matematica tutta ad altrettanta sicurezza attenendosi al «modo finitista» di pensare, una volta che si sposti l'attenzione dagli oggetti alle formule che degli oggetti parlano. E ciò si ottiene definitivamente con un secondo principio, in base al quale le formule che consideriamo non hanno in sé alcun significato anche se ovviamente sono poste come significanti qualcosa. Essi servono solo a comunicare asserzioni. Prendiamo ad esempio il celebre teorema di Euclide che afferma che fra p, intero primo, e $p! + 1$ esiste certamente un numero primo. Tale teorema è *finitista* perché la sua enunciazione può essere evidentemente tradotta nella disgiunzione *finita* «$p + 1$ è primo, oppure $p + 2$ è primo, ..., oppure $p! + 1$ è primo». Orbene si fa paradossalmente un «salto» nel transfinito dando una formulazione *meno forte* di questo teorema, dicendo ad esempio semplicemente: esiste un numero primo maggiore di p; in questo caso infatti la disgiunzione precedente si trasforma in una disgiunzione con un numero *infinito* di termini. Dal punto di vista finitista quindi un'espressione esistenziale della for-

ma $\exists xPx$ non è interpretabile; analogamente ci immettiamo nel transfinito negando una proposizione universale, e quindi, dal punto di vista finitista, questa non è un'operazione che ci porta ad enunciati interpretabili: nel transfinito non valgono le leggi della logica aristotelica.

Ora Hilbert vuole conservare quello che è l'usuale procedere matematico, nel quale appunto si fanno ordinariamente affermazioni esistenziali e si negano proposizioni universali; e vuole allora risolvere le difficoltà sopra prospettate proprio ricordando di essere un *matematico*. Sicché facendo esplicito riferimento alla procedura usata in altre parti della matematica,[44] egli propone di riconoscere, accanto alle proposizioni interpretabili direttamente, in termini concreti, finiti (contenutistiche, materiali o *reali*) delle *proposizioni ideali* per «mantenere le regole dell'ordinaria logica aristotelica, formalmente semplici». La matematica diventa così «un inventario di formule, in primo luogo formule alle quali corrispondono le comunicazioni contenutistiche delle proposizioni finitarie (quindi, in massima parte, uguaglianze e disuguaglianze numeriche) e in secondo luogo altre formule che in se stesse non significano niente e che sono *gli oggetti ideali della nostra teoria*».

Già il diverso atteggiamento nei riguardi della portata e della importanza della logica classica segna una netta differenziazione con l'atteggiamento intuizionista; in particolare, afferma Hilbert, il terzo escluso «è conseguenza dell'assioma logico sull'ε [si veda più avanti] e non ha mai causato il minimo errore»: privare il matematico di questo principio sarebbe come togliere il telescopio all'astronomo. Ma ancora, all'affermazione di Brouwer, il quale «dichiara (proprio come Kronecker) che le proposizioni esistenziali sono in sé prive di significato a meno che esse non contengano anche la costruzione dell'oggetto di cui si asserisce l'esistenza; ... il loro impiego, per lui, fa degenerare la matematica in un gioco», corrisponde una chiarissima enunciazione del significato profondo di questi enunciati a livello sistematico, la consapevolezza che – pena l'amputazione della matematica – occorre fare «come se» esse avessero sempre un significato. Dice Hilbert: «il valore delle dimostrazioni puramente esistenziali consiste precisamente

[44] Ad esempio in algebra, con l'introduzione dei «numeri ideali» o in geometria proiettiva, con l'introduzione dei «punti impropri» o all'infinito.

nel fatto che loro tramite si possono eliminare costruzioni individuali e molte costruzioni differenti sono sussunte sotto un'idea fondamentale, cosicché risulta chiaramente solo ciò che è essenziale alla dimostrazione; brevità ed economia di pensiero sono *la raison d'être* delle dimostrazioni d'esistenza».

Questo intergioco tra problemi, costruzioni che permettono di risolverli, assunzioni che garantiscono la possibilità delle costruzioni, è per Hilbert il lievito dello sviluppo matematico. Il matematico parte da problemi (si ricordi la famosa conferenza del 1900 a Parigi) cerca di risolverli con i mezzi (gli assiomi) dati e, se non vi riesce, si garantisce nuove possibilità di risoluzione introducendo nuovi assiomi. In questo modo si amplia il dominio degli oggetti considerati e accanto alle proposizioni materiali si pongono proposizioni ideali che coinvolgono enti astratti, il cui ruolo è quello di permettere la risoluzione di interrogativi sugli oggetti precedentemente ammessi. In questa prospettiva due sono i problemi da affrontare: il primo è quello di avere un metodo per decidere quanto un dato problema è risolubile con dati mezzi (è questo l'*Entscheidungsproblem* di Hilbert nella sua forma più generale), il secondo è quello di garantire la liceità dell'assunzione di assiomi ed è a questo punto che interviene la *Beweistheorie* e si pone il problema della coerenza.

L'organizzazione concettuale della matematica esige il ricorso all'infinito e alle proposizioni ideali ma non è necessario che singolarmente ogni enunciato sia interpretabile finitisticamente. È il sistema assiomatico nel suo complesso che va analizzato e giustificato. Il problema è allora quello di indagare a che condizione i sistemi assiomatici sono accettabili da un punto di vista finitista. Come abbiamo detto è la coerenza che ci dice quando una teoria assiomatica può essere adottata e si tratta allora di provare la coerenza delle singole teorie senza presupporre nulla che rimandi a enti ideali e quindi senza far ricorso alla costruzione di modelli e metodi analoghi, che forniscono solo dimostrazioni *relative* di coerenza. Dimostrazioni *assolute* le possiamo ottenere solo utilizzando i metodi intuitivamente giustificati che appartengono al pensiero finitista. L'obiettivo del programma hilbertiano diviene così quello di provare finitisticamente la coerenza delle teorie matematiche. Ciò avviene in due passi.

Il primo passo ci porta dalla teoria quale si presenta nella prati-

ca e in cui si fanno riferimenti infinitari a oggetti non concreti, a teorie formalizzate entro specifici linguaggi in cui i simboli primitivi, le regole di formazione e le regole di dimostrazione devono essere specificate in modo dominabile. Esempi di linguaggi e di calcoli di questo tipo sono quelli che abbiamo già visto parlando della frammentazione della Grande Logica.

Il secondo passo consiste nel prendere la teoria formalizzata come oggetto di indagine metamatematica e di provare con mezzi puramente finitisti che per la teoria in esame non è possibile dimostrare enunciati del tipo $\mathscr{A} \wedge \neg \mathscr{A}$. Questo comporta un'analisi dettagliata della struttura delle dimostrazioni formali che – diversamente dagli oggetti di cui parla la teoria – sono oggetti concreti, successioni di simboli: è questa la ragione per cui lo strumento del programma hilbertiano è costituito dalla teoria della dimostrazione.

Nella costruzione di tale teoria Hilbert fa uso di una particolare versione dei linguaggi del primo ordine, che si basa sulla introduzione di uno speciale operatore, l'operatore ε, in termini del quale risultano definibili i quantificatori. Nelle intenzioni di Hilbert, questo operatore concentra in sé la somma dei riferimenti infinitari che si fanno utilizzando i quantificatori ed è strettamente legato ad un'interpretazione logica dell'assioma di scelta. L'assioma fondamentale che riguarda l'ε-operatore è così il seguente

$$(1) \qquad \mathscr{A}(t) \rightarrow \mathscr{A}(\varepsilon x \mathscr{A})$$ [45]

che possiamo leggere come: dato un qualunque individuo t che gode di $\mathscr{A}$, allora anche l'oggetto $\varepsilon x \mathscr{A}$ godrà di $\mathscr{A}$. In questo modo l'applicazione di ε sceglie fra i possibili oggetti che godono di $\mathscr{A}$ un oggetto generico nel senso che, come fissato dall'assioma (1), esso costituisce una sorta di esemplificazione ambigua della proprietà. Non stupisce allora che, se adottiamo le regole della logica del primo ordine già specificate, otterremo che i quantificatori risultano caratterizzabili in termini dell'ε-operatore, nel senso che valgono

$$\exists x \mathscr{A} \leftrightarrow \mathscr{A}(\varepsilon x \mathscr{A})$$
$$\forall x \mathscr{A} \leftrightarrow \mathscr{A}(\varepsilon x \neg \mathscr{A}).$$

[45] Questo assioma, dice Hilbert, «contiene il nucleo di uno dei più controversi assiomi della letteratura matematica, ossia l'assioma di scelta».

Introdotto nei primi anni venti da Hilbert, l'ε-operatore fu utilizzato nel 1924 nella sua tesi da Ackermann e studiato sistematicamente da Hilbert e Bernays nella loro opera del 1939. Qui venivano per la prima volta formulati due teoremi fondamentali relativi ad esso, i cosiddetti ε-teoremi, collegati ai teoremi di Herbrand e di Gentzen di cui parleremo più avanti. Entrambi i risultati fanno riferimento ad un calcolo **F** (l'ε-calcolo) che si ottiene dall'usuale calcolo del primo ordine aggiungendo l'assioma (1). Il primo ε-teorema afferma allora che se $\mathscr{A}$ è una formula derivabile in **F** da $\mathscr{A}_1, \ldots, \mathscr{A}_k$ e se tanto $\mathscr{A}$ quanto $\mathscr{A}_1, \ldots, \mathscr{A}_k$ non contengono variabili vincolate, allora $\mathscr{A}$ può essere derivata da $\mathscr{A}_1, \ldots, \mathscr{A}_k$ utilizzando le sole regole del calcolo proposizionale. Il secondo ε-teorema afferma invece che se $\mathscr{A}$ è una formula che non contiene l'ε-simbolo dimostrabile in **F** da $\mathscr{A}_1, \ldots, \mathscr{A}_k$, allora $\mathscr{A}$ è dimostrabile all'interno del calcolo dei predicati senza l'ε-simbolo.

In entrambi i teoremi si assume che l'ε-simbolo non compaia negli assiomi extralogici $\mathscr{A}_1, \ldots, \mathscr{A}_k$. Sarà il calcolo **F** a costituire lo strumento principe della teoria della dimostrazione sino agli interventi di Herbrand e di Gentzen e i due ε-teoremi avranno un ruolo centrale in quanto da essi si possono ottenere condizioni sulla non contraddittorietà delle teorie.

A questo proposito c'è un'osservazione da fare, che mostra un aspetto più profondo del concetto di coerenza, che Hilbert poneva al centro dell'attenzione. Come si è detto, l'introduzione degli elementi ideali è soggetta ad una sola condizione, quella della coerenza (non contraddittorietà) nel senso che la loro aggiunzione al vecchio dominio non deve comportare contraddizioni. La cosa ha un significato ben preciso. Infatti dal principio di contraddizione che Hilbert assume nella forma $(\mathscr{A} \rightarrow (\mathscr{B} \wedge \neg \mathscr{B})) \rightarrow \neg \mathscr{A}$ si ricava la formula $(\mathscr{A} \rightarrow \neg \mathscr{A}) \rightarrow \mathscr{B}$, vale a dire, sulla base della regola di separazione: se nella teoria è derivabile una contraddizione, ossia una espressione della forma $\mathscr{A} \wedge \neg \mathscr{A}$ per una qualche $\mathscr{A}$, allora è possibile derivare nella teoria *qualunque* formula $\mathscr{B}$, ad esempio la formula $1 \neq 1$. Dimostrare allora la coerenza del sistema significherà mostrare che $1 \neq 1$ non è una formula dimostrabile. Poiché $1 \neq 1$ è una proposizione contenutistica falsa, possiamo dire in altro modo che la condizione di non contraddittorietà comporta (e di fatto, si può provare, equivale) al fatto che l'aggiunzione di proposizioni ideali è *conservativa* rispetto a quelle reali,

cioè non ci porta a provare proposizioni contenutistiche che non fossero già dimostrabili finitisticamente senza di esse. È questo il legame profondo tra noncontraddittorietà e aggiunzione di principi ideali.

Il campo d'azione della teoria della dimostrazione non si limita però al problema della coerenza, ma ci permette di affrontare anche l'*Entscheidungsproblem*, offrendo a Hilbert l'occasione di affermare la risolubilità di ogni problema. Scrive Hilbert: «come esempio del modo con cui si possono trattare le questioni fondazionali, voglio scegliere la tesi che ogni problema può essere risolto... Infatti, in matematica non esistono *ignorabimus*. Ora la mia teoria della dimostrazione non può specificare un metodo generale per risolvere ogni problema matematico; un tale metodo non esiste. Ma la dimostrazione che l'assunzione della risolubilità di ogni problema matematico è coerente rientra nell'ambito della nostra teoria». Questa ipotizzata possibilità di risolvere tutti i problemi matematici, nella quale sarà facile vedere da parte intuizionista un'affermazione ulteriore del principio del terzo escluso, sarà un altro punto di insormontabile attrito tra i due indirizzi, come presto vedremo. In conclusione Hilbert si ritiene già in questo periodo in condizione di poter anticipare quello che sarà l'«esito finale» delle ricerche sui fondamenti: «la matematica è una scienza senza ipotesi. Per provarlo non ho bisogno di dio, come fa Kronecker, o dell'assunzione di una speciale capacità del nostro intelletto relativa al principio di induzione matematica come fa Poincaré, o dell'intuizione originaria di Brouwer o, infine, come fanno Russell e Whitehead, degli assiomi dell'infinito, di riducibilità, o di completezza, che in effetti sono assunzioni contenutistiche che non possono essere giustificate con una dimostrazione di coerenza».

I primi risultati non tardarono ad arrivare e la *Beweistheorie* sembrò fornire le desiderate conferme quando Ackermann, sviluppando la teoria dell'operatore ε ritenne di aver dimostrato per suo mezzo in *Begründung des* tertium non datur *mittels den Hilbertschen Theorie der Widerspruchfreiheit* (*Fondazione del* tertium non datur *per mezzo della teoria hilbertiana della noncontraddittorietà*, 1927) la coerenza dell'Analisi; successivamente si accorse di aver commesso un errore nella dimostrazione e il risultato venne ridimensionato nello stesso anno da von Neumann nell'articolo *Zum Hilbertschen Be-*

weistheorie (*Per la teoria della dimostrazione di Hilbert*), nel quale dava una dimostrazione di noncontraddittorietà per un sistema ristretto, ossia l'aritmetica senza induzione, o meglio con l'induzione limitata a proprietà particolari, la cui espressione linguistica non contenga quantificatori (proprietà, in altri termini, che non si riferiscono alla totalità dei numeri naturali). Di questo risultato, nel 1929 J. Herbrand darà una dimostrazione semplice utilizzando tecniche di sua invenzione, che non dipendono dall'ε-calcolo hilbertiano e su cui torneremo parlando della completezza. Ci basti dire che queste indagini culmineranno nel 1930 e nel 1931 con le *Recherches sur la théorie de la démonstration* (*Ricerche sulla teoria della dimostrazione*), sua tesi di laurea, e *Sur le problème fondamental des mathématiques* (*Sul problema fondamentale delle matematiche*) dove Herbrand svilupperà la sua versione della teoria della dimostrazione formulando il suo teorema fondamentale e applicandolo a diverse questioni, in particolare a casi specifici dell'*Entscheidungsproblem* hilbertiano. Rimandiamo a paragrafi successivi l'analisi di alcuni di questi risultati. Quello che ci preme sottolineare è che al sorgere degli anni trenta il progetto di una metamatematica su basi finitiste non era più solo un'idea ma una realtà in sviluppo, con tecniche diverse e un piccolo nucleo di risultati centrali ed importanti.

Già sulla base di questi risultati parziali i formalisti – ma sostanzialmente il mondo matematico in generale – pur con le dovute cautele potevano nutrire serie speranze nell'esito finale della loro impresa, come mostrano chiaramente le seguenti parole di von Neumann, scritte nel 1930: «Malgrado la coerenza della matematica classica non sia ancora stata dimostrata, una tale dimostrazione è stata trovata per un sistema matematico alquanto più ristretto, che è strettamente correlato a un sistema che Weyl aveva proposto prima della concezione del sistema intuizionista, ma più ristretto della matematica classica. Così il sistema di Hilbert ha superato il primo test di efficacia: è stata stabilita con metodi finitari costruttivi la validità di un sistema matematico non finitario, non puramente costruttivo. Se qualcuno riuscirà ad estendere questa garanzia al sistema più impegnativo e importante della matematica classica, potrà dirlo solo il futuro.»

Il riferimento a Weyl non era casuale. Già nel 1918 in *Das Kontinuum, kritische Untersuchungen über die Grundlagen der Analysis* (*Il continuo. Ricerche critiche sui fondamenti dell'Analisi*) egli si era decisa-

mente pronunciato per una «restrizione» dell'Analisi, avendo riconosciuto insuperabile la difficoltà legata alla predicatività; Weyl qualificava infatti l'Analisi impredicativa (nella quale cioè siano ammessi procedimenti o definizioni impredicative) «come una casa costruita sulla sabbia» in una sorta di «paradiso dei logici» privo di contenuto matematico reale. Weyl è convinto che il fondamento ultimo del pensiero matematico sia «la rappresentazione dell'iterazione, della successione dei numeri naturali» e identificata l'eccellenza della matematica stessa nel fatto «che in quasi tutti i suoi teoremi ciò che è per sua natura *infinito* viene ricondotto a una decisione finita», ribadisce che «questa infinità dei problemi matematici riposa però sul fatto che la *successione dei numeri naturali e il concetto di esistenza che ad essa si riferisce* ne costituiscono la base». Una posizione decisamente orientata verso il neointuizionismo brouweriano (come non tardò a realizzare Hilbert con estremo rincrescimento), che tuttavia si sviluppava in modo assai originale, con la creazione di due sistemi di *Analisi predicativa* che non è però possibile qui descrivere compiutamente. Ci limitiamo a osservare – con le parole di Casari – che intuitivamente questi due sistemi[46] vengono edificati a partire da un «dominio operativo» costituito da un certo numero di categorie fondamentali di enti, per i quali sono date determinate proprietà e determinate relazioni primitive. A partire da questi dati si costituisce la teoria tramite: *un processo logico* «che consiste nel generare, a partire da un dato stock iniziale di proprietà e relazioni *primitive* che si riferiscono agli enti di determinate *categorie iniziali*, tutta una gamma di nuove proprietà e relazioni riferite agli stessi enti» e un *processo matematico* attraverso il quale «si può passare – a partire da un certo insieme di proprietà e relazioni, riferite a certi enti – alla costituzione di nuove *entità ideali*».

Nel 1927 tuttavia la posizione di Weyl muta e in *Diskussionbemerkungen zu dem zweiten Hilbertschen Vortrag* [1927] *über die Grundlagen der Mathematik* (*Commenti sulla seconda conferenza di Hilbert sui fondamenti della matematica*) diviene più mediata nel senso che pur difendendo Brouwer da alcune critiche di Hilbert, egli mette in evidenza nel contempo la portata e il significato dell'approccio

[46] Il primo dei quali è «sostanzialmente equivalente alla teoria ramificata dei tipi senza assioma di riducibilità e l'altro assai più debole» (Casari).

hilbertiano. In particolare, 1) si oppone ad Hilbert (in favore dell'intuizionismo, ma sostanzialmente di Poincaré) sulla questione della giustificazione dell'induzione matematica; non stima cioè soddisfacente la risposta di Hilbert a Poincaré, malgrado veda parecchi punti di convergenza fra i due; 2) concorda con Brouwer nel concepire la matematica come costituita da «proposizioni reali» e ritiene che egli abbia visto per primo «esattamente e in tutta la sua portata come in effetti ci si era ovunque allontanati dai limiti del pensiero contenutistico»; 3) concorda tuttavia con Hilbert nell'ammettere la possibilità di trattare con proposizioni ideali tramite le quali la matematica va al di là «di stati di cose intuitivamente accertabili». Ci sembra che le seguenti parole conclusive di Weyl ben rispecchino e l'atteggiamento generale di larga parte dei matematici e il suo proprio: «Se le vedute di Hilbert prevarranno sull'intuizionismo, come sembra in effetti debba avvenire, *allora vedo in ciò una sconfitta decisiva dell'attitudine filosofica della fenomenologia*, che così si dimostra insufficiente per la comprensione della scienza creativa anche nell'area della conoscenza che è la più primaria e la più aperta all'evidenza, la matematica» dove, si badi bene, le simpatie di Weyl sono tutte per la fenomenologia e non certo per il formalismo hilbertiano.

Ma altre posizioni di ispirazione finitista facevano la loro comparsa in questo periodo, prima fra tutte quella del norvegese Thoralf Skolem. Qui vogliamo considerare un suo lavoro dal titolo *Begründung der elementaren Arithmetik durch die rekurrierende Denkeweise ohne Anwendung scheinbarer Verändlichen mit unendlichen Ausdehnungsbereich* (*Fondazione dell'aritmetica elementare per mezzo del modo ricorsivo di pensare, senza l'impiego di variabili apparenti varianti su domini infiniti*), pubblicato nel 1923, ma scritto per sua esplicita dichiarazione già nel 1919, «dopo aver studiato l'opera di Russell e Whitehead», ossia i *Principia*.

In questo articolo Skolem presentava un nuovo modo di sviluppare l'aritmetica eliminando in certo senso alla radice la nozione stessa che produceva i paradossi ed era responsabile della complessità della teoria ramificata dei tipi intesa ad evitarli: la nozione di totalità (attuale) infinita, o, linguisticamente, l'impiego dei quantificatori illimitati per variabili individuali.

Al posto di questo riferimento alla totalità dei naturali via

quantificatori, Skolem sostituiva l'uso sistematico delle procedure induttive e ricorsive. Ecco come lo stesso Skolem presentava il suo piano di lavoro: «Se consideriamo i teoremi generali dell'aritmetica come asserzioni funzionali e assumiamo come base il modo ricorsivo di pensare, allora questa scienza può essere fondata in modo rigoroso senza l'uso delle nozioni "sempre" [tutti] e "qualche volta" [almeno uno]... si può assicurare una fondazione logica per l'aritmetica senza l'uso di variabili logiche apparenti [cioè senza variabili quantificate].

In verità sarà spesso vantaggioso introdurre variabili apparenti, ma richiederemo che queste variabili varino solo su domini finiti, e per mezzo di definizioni ricorsive saremo allora sempre in grado di evitare l'uso di tali variabili». Naturalmente l'impiego delle sole variabili libere o «reali» come le abbiamo a suo tempo chiamate (o di variabili apparenti vincolate da quantificatori limitati) diminuisce di molto la «potenza espressiva» del linguaggio impiegato (che noi possiamo assumere essere un linguaggio elementare per l'aritmetica). Ad esempio si può esprimere e dimostrare che esistono infiniti numeri primi servendoci del teorema di Euclide ricordato qualche pagina addietro perché esso afferma per ogni numero primo dato l'esistenza di almeno un altro numero primo *entro limiti numerici precisati*; ma non potremo ad esempio esprimere che esistono infinite coppie di primi gemelli ossia primi della forma $(p, p + 2)$ (ad esempio, (5, 7), (11, 13) ecc.).

Orbene Skolem mostra come questa limitazione si possa in parte neutralizzare impiegando *schemi ricorsivi* per l'introduzione di nuove funzioni o relazioni e impiegando nelle dimostrazioni l'induzione matematica *come regola*, ossia ammettendo conclusioni della forma $P(x)$ ogniqualvolta si siano verificate le premesse $P(0)$ e $P(x) \rightarrow P(x')$.[47]

Appunto questo è ciò che Skolem chiama il «modo ricorsivo di

[47] Un esempio di definizione ricorsiva è quello di somma dato dalle equazioni $x + 0 = x$; $(x + y') = (x + y)'$. Ancora, la relazione «$<$» fra numeri naturali viene di solito definita come segue: $x < y \leftrightarrow \exists z(x + z = y)$; il quantificatore può essere eliminato mediante la definizione ricorsiva; $x < 0$ è falso; $x < y' \leftrightarrow (x < y) \vee (x = y)$. Come si vede qui la funzione o la relazione sono definite dando il loro valore per 0 e quindi calcolando il valore per ogni numero successivo sulla base di valori precedentemente ottenuti. Nel capitolo V, paragrafo 6, vedremo come le considerazioni di Skolem siano alla base della teoria della ricorsività nata negli anni trenta.

pensare» ed il sistema basato su questi principi è grosso modo quello che oggi noi diciamo *aritmetica ricorsiva primitiva* (ARP). La verificabilità e costruttività intrinseca del metodo portavano Skolem (e lui stesso lo dichiara, anche se ribadisce l'indipendenza) assai vicino all'intuizionismo; in effetti tuttavia l'aritmetica primitiva ricorsiva è ancor più ristretta dell'aritmetica intuizionista e il suo accentuato carattere finitista la porta più precisamente nell'ambito di una concezione hilbertiana. Non a caso Hilbert e Bernays che nel I volume delle loro *Grundlagen* dettero un'ampia esposizione delle idee di Skolem, riconoscono nell'ARP una formulazione adeguata del nucleo finitista dell'aritmetica vedendo in essa il contesto appropriato per la formalizzazione del discorso *metamatematico* della teoria della dimostrazione. Ma avremo occasione di interessarci più da vicino dell'argomento nel paragrafo 6 del capitolo V. Qui volgiamo solo ricordare che l'idea di sviluppare non solo l'aritmetica ma anche l'Analisi su basi ricorsive fu ripresa e sviluppata da Richard Goodstein (1912-85) negli anni cinquanta.

3.5 *Brouwer e la scuola intuizionista*

Abbiamo già delineato le posizioni di Brouwer e il suo intervento nei primi anni del secolo nelle discussioni sui fondamenti della matematica. Negli anni venti la produzione fondazionale di Brouwer diviene copiosissima ed è diretta da una parte alla presentazione di tutta una serie di concetti e risultati della matematica intuizionista, dall'altra a una vigorosa polemica contro la concezione hilbertiana (abbiamo visto come il logicismo fosse stato ormai, in certo senso, «sconfitto»). Rimandiamo al prossimo capitolo un'esposizione più dettagliata della matematica intuizionista che man mano si andava sviluppando per opera di Brouwer e dei collaboratori. Qui vogliamo limitarci ad analizzare un elemento importante nella polemica. Ci riferiamo al primo tentativo di formalizzazione del calcolo logico intuizionista avvenuto nel 1925 ad opera di Andrei Nikolaevich Kolmogorov (1903-88) nell'articolo *Sul principio del "tertium non datur"*. Va subito detto che Brouwer guardò con sospetto a questo primo tentativo e non cambiò atteggiamento neppure riguardo ai successivi lavori in questo senso di Heyting degli anni trenta e che noi vedremo in un prossimo capitolo.

Si ricorderà che il momento fondamentale del pensiero di Brouwer, che fa sì che la sua posizione sia tanto originale e «scomoda», nel contesto delle ricerche sui fondamenti è la tesi che oggetto della matematica sono degli enti che richiedono una forma particolare di logica: le costruzioni (matematiche) mentali. La matematica si identifica con la parte esatta del nostro pensiero e la sua unica fonte è l'intuizione, da non intendersi quest'ultima in senso metafisico o mistico, bensì come capacità di considerare separatamente determinati concetti e costruzioni che intervengono normalmente nel nostro pensare abituale. In particolare, ogni teorema logico non è che un teorema matematico di estrema generalità: vale a dire, la logica è una parte della matematica e non può quindi in alcun modo servire per la fondazione di quest'ultima. Questa convinzione di fondo, che ha riflessi immediati sulla concezione brouweriana dell'esistenza degli enti matematici,[48] comportava inoltre tutta una serie di prese di posizioni estremamente radicali, relative alla logica e al linguaggio, che da una parte rendono la concezione intuizionista estremamente suggestiva, dall'altra tutto sommato difficilmente confutabile.

Per quanto riguarda il linguaggio, secondo Brouwer, la matematica ne è in linea di principio indipendente, e questo per il semplice motivo che il pensiero e le costruzioni mentali non hanno alcun bisogno di essere connessi con espressioni linguistiche. Ovviamente il linguaggio mantiene il suo valore di *comunicazione* delle esperienze matematiche dei diversi ricercatori, ma quel che in definitiva conta è proprio il *sistema matematico* che ogni singolo ricercatore edifica nella sua mente. Nell'atto della comunicazione linguistica agli altri, nessun linguaggio può garantire da malintesi, fraintendimenti od errori intrinsecamente connessi a ogni forma di espressione linguistica. In altri termini, *non esiste un linguaggio non equivoco per la matematica*. Su questo tema della separazione netta fra matematica e linguaggio, che è peculiare della concezione brouweriana, ecco cosa lo stesso Brouwer dice, in una conferenza tenuta nel marzo 1928 dal titolo *Mathematik, Wissenschaft und Sprache* (*Matematica, scienza e linguaggio*): «*Non esiste quindi anche per la*

[48] Si ricordi che il programma brouweriano riguarda proprio l'indagine delle costruzioni matematiche mentali in quanto tali, senza riferimento alle questioni relative alla natura degli oggetti costruiti, come se questi oggetti esistessero indipendentemente da noi.

matematica pura nessun linguaggio sicuro, ossia nessun linguaggio che escluda il presentarsi di fraintendimenti e che con l'accuratezza assicuri la libertà dall'errore (ossia dello scambio di entità matematiche diverse). Né si può migliorare questa circostanza sottoponendo, come fa *la scuola formalista*, a una considerazione matematica la stessa lingua matematica (cioè il sistema di segni che serve ad altri uomini per nominare costruzioni di matematica pura), attribuendole con manipolazioni l'adeguatezza e la stabilità di uno strumento materiale o di un fenomeno della scienza esatta e si riferisca di essa in una lingua di secondo ordine o superlingua. E infatti, in primo luogo questa superlingua (poiché si riferisce a un insieme finito e dominabile di oggetti concreti e alla matematica pura di un sistema finito da essi astratta) può certo, grazie all'uso della lingua matematica, mettere al riparo da errori o fraintendimenti: ma proprio per l'essenza stessa della lingua non con sicurezza assoluta; in secondo luogo, anche se si verificasse appunto la certezza assoluta non sarebbe per nulla eliminata la possibilità di fraintendimenti nei riguardi delle pure costruzioni matematiche alluse attraverso una tale lingua matematica esatta.»

In altri termini, l'intuizionista non è interessato al lato linguistico-formale della matematica, ma soltanto a quel tipo di ragionamento che appare in matematica. Ne risulta immediatamente, come vedremo meglio, l'impossibilità di una convergenza fra formalismo e intuizionismo sulla base di una formalizzazione della matematica intuizionista. Se infatti si intende la formalizzazione come un linguaggio per la matematica, non si eliminano, per la natura stessa generale del linguaggio, possibilità di errori o malintesi; se invece per formalizzazione si intende una struttura matematica particolarmente semplice, allora sorge il pericolo che il sistema formale non sia in grado di rappresentare pienamente il dominio della matematica, con la conseguente necessità di ampliare il sistema ogniqualvolta si è messi di fronte a una nuova scoperta.

La questione ci porta direttamente anche all'altro aspetto caratteristico della posizione intuizionista, che riguarda appunto il rapporto tra logica e matematica. Infatti a parere di Brouwer la fiducia dei formalisti nei poteri «magici» del linguaggio deriva dalla loro convinzione circa la completa e indubitabile validità della *logica classica*. Tale fede ha un'origine molto antica, sulla base della

quale sono state storicamente messe in evidenza alcune forme privilegiate di passaggi da proposizioni a proposizioni, che sono poi state cristallizzate nei cosiddetti principi logici, quali ad esempio il principio di identità, il principio di contraddizione, il principio del terzo escluso e il principio del sillogismo. Ora questi principi assolvono pienamente la loro funzione se impiegati a livello linguistico ordinario o anche in un linguaggio scientifico le cui asserzioni siano tuttavia «controllabili». Ma le difficoltà sorgono quando si passa al caso nel quale le proposizioni ottenute tramite tali principi vengono considerate come acquisite proprio solo *in forza dei principi stessi*, ossia anche quando esse non siano suscettibili di essere vagliate attraverso un controllo *diretto*. Questo in particolare avviene proprio nel caso del principio cruciale per Brouwer, il principio del terzo escluso, al quale si accorda questa fiducia anche «... nella forma estesa secondo la quale un evento precedente viene assunto come avvenuto non soltanto sulla base dell'assurdità, ma anche semplicemente dell'impossibilità pratica di trovare un'altra spiegazione per un fatto da stabilirsi». In definitiva, la fiducia accordata ai principi logici, e in particolare al principio del terzo escluso, trae in effetti la sua ragion d'essere dal carattere essenzialmente *finito* dell'organizzazione della conoscenza umana, e trova soltanto in ciò la sua giustificazione ultima. La logica umana è *finita*, i suoi principi valgono senza limitazioni solo nel caso in cui si applichino a sistemi di qualunque tipo ma *finiti*: per poter attuare un'estensione di tali principi al caso infinito si rende necessaria una verifica della loro validità; e questo riesame mostra in particolare che il principio del terzo escluso non presenta questa validità universale. Accettare quindi in generale tali principi senza alcuna verifica, estendere con leggerezza la loro portata dal finito all'infinito comporta le contraddizioni ben note (che secondo Brouwer non possono essere eliminate mediante una semplice «rielaborazione» degli assiomi) e l'introduzione nella matematica di «proposizioni ideali» che dal punto di vista della concezione contenutistica brouweriana sono semplicemente null'altro che «parole vuote».

In definitiva, per quanto riguarda il rapporto della logica con la matematica, dal punto di vista intuizionista le conclusioni di un ragionamento non derivano dall'impiego di schemi logici prefissati e cristallizzati, ma si impongono, per così dire, per la loro stessa

evidenza (è chiaro che gli schemi logici possono, peraltro, avere una loro importanza per *sistematizzare* la materia, cosa però questa che Brouwer afferma con una certa reticenza). Ne risulta che da un lato la matematica è del tutto indipendente dalla logica, dall'altro che la logica matematica è null'altro che una matematica applicata.

Tra i principi logici che hanno un significato solo formale fuori dal dominio del finito – come si ricorderà – il terzo escluso è il più pericoloso: è in base ad esso che vengono condotte le dimostrazioni per assurdo che tipicamente ci portano a conclusioni non realizzabili con costruzioni: per Hilbert il terzo escluso era come «i guantoni per il pugile», uno strumento indispensabile e rinunciarci significava abbandonare l'idea stessa che ogni problema in qualche modo è risolubile e quindi il progetto stesso della ricerca scientifica. Per Brouwer non è così e per capire in che senso egli negasse validità al terzo escluso, partiamo da un esempio, che prendiamo da Heyting. Consideriamo le due definizioni:

1) k è il massimo numero primo tale che anche $k - 1$ è primo, oppure $k = 1$ se tale numero non esiste;
2) l è il massimo numero primo tale che anche $l - 2$ è primo, oppure $l = 1$ se tale numero non esiste.

Si vede subito che 1) è una definizione «costruttiva», ossia permette di individuare effettivamente il numero che definisce (in questo caso $k = 3$), mentre per l'intuizionista 2) non è una definizione (*non* individua cioè il numero l), perché non sappiamo se il numero delle coppie di numeri primi gemelli è finito o no; se tuttavia si ammette il terzo escluso (come appunto fa il matematico «classico») la successione di primi gemelli $(p, p + 2)$ risulta necessariamente finita o infinita e quindi anche la 2) definisce senza ambiguità un numero (indipendentemente dal fatto che noi sappiamo dire o no di quale numero si tratti).

Se la matematica è un'attività mentale che per quanto connessa ad operazioni formali non può assolutamente essere ridotta ad esse, gli oggetti matematici saranno soltanto costruzioni mentali che il soggetto riesce a compiere partendo da materiali intuitivi. Due – abbiamo visto – sono queste intuizioni: noi sappiamo che possiamo pensare a entità separate le une dalle altre e sappiamo che siamo capaci di pensare successioni indefinitamente

protraentesi di tali entità. Queste nozioni sono *immediatamente presenti* alla nostra mente, non abbisognano di alcuna spiegazione o fondazione; esse sono la «sorgente» del concetto di numero naturale e «la matematica comincia dopo che sono stati costituiti i concetti di numero naturale e di uguaglianza fra numeri naturali». L'assioma di induzione completa va riguardato come un teorema generale (di immediata e semplice «dimostrazione» intuizionista) sui numeri naturali, e tutti gli altri assiomi di Peano sono semplicemente evidenti. Non c'è alcun problema di fondazione come pretendevano logicisti e formalisti: si tratta solo di intendere i numeri in base alla intuizione della biunità. È quindi questa nozione elementare di numero naturale che l'intuizionista assume come fondamentale per la sua matematica, considerandola come una nozione familiare e nota a ogni «creatura pensante», sulla base della quale effettuare costruzioni (matematiche) mentali. Per tale nozione gli intuizionisti non reclamano «alcuna forma di certezza o di definitezza in senso assoluto, che sarebbe irrealizzabile», ma sostengono che essa «è sufficientemente chiara per costruirvi sopra la matematica». In particolare allora un teorema matematico viene a esprimere un fatto puramente empirico: esso afferma che è stata effettuata una certa costruzione. Rimandiamo a un prossimo capitolo una presentazione di come concretamente si sviluppa la matematica intuizionista e prima di passare ai risultati centrali del citato lavoro di Kolmogorov, per quanto in particolare riguarda il rapporto intuizionismo/formalismo (che ha tutto sommato in questi anni tutto l'aspetto di un «dialogo» fra sordi)[49] vogliamo ricordare che nel 1928 Brouwer nell'articolo *Intuitionistische Betrachtungen über den Formalismus* (*Considerazioni intuizioniste sul formalismo*) proponeva quattro punti che, se accettati, avrebbero posto fine alla diatriba sui fondamenti e avrebbero reso il propendere per l'intuizionismo o per il formalismo solo «una questione di gusto». Mette conto di riportare – con minime modificazioni rispetto al

[49] Dice Heyting, presentando i due come avversari irriducibili: «Come Brouwer non poteva sopportare di vedere la matematica, questo gioiello prezioso dello spirito, divenire un gioco di segni privo di significato, analogamente era impossibile per Hilbert sacrificare le più belle parti di questo maestoso edificio a una mania di sottigliezza». In effetti non poche discussioni erano sterili in quanto a termini linguistici uguali venivano dati spesso, dalle due parti, significati radicalmente diversi.

testo brouweriano – i quattro punti in questione («intuizioni» li chiama Brouwer).

1) «La differenziazione, nei lavori formalisti, fra una costruzione dell'"inventario delle formule matematiche" (visione formalistica della matematica) e una teoria intuitiva (contenutistica) delle leggi di questa costruzione, col riconoscimento inoltre che per quest'ultima teoria è indispensabile la matematica intuizionista dell'insieme dei numeri naturali.»

2) «Il rifiuto dell'applicazione meccanica del principio del terzo escluso e il riconoscimento: 1) del fatto che l'indagine delle basi giustificative del principio in questione e del suo ambito di validità costituisce un argomento essenziale delle ricerche sui fondamenti della matematica; 2) del fatto che nella matematica intuitiva (contenutistica) questo principio è valido solo per sistemi finiti».

3) «L'identificazione del principio del terzo escluso col principio della risolubilità di ogni problema matematico.»

4) «Il riconoscimento del fatto che la giustificazione (contenutistica) della matematica formalistica per mezzo della dimostrazione di coerenza contiene un circolo vizioso, poiché questa giustificazione si fonda sulla correttezza (contenutistica) della proposizione che dalla coerenza di una proposizione segua la correttezza della stessa, ossia sulla correttezza (contenutistica) del principio del terzo escluso».

Presentate le sue «intuizioni» Brouwer mostrava, con precise citazioni testuali, che le prime due erano sostanzialmente accettate nella letteratura formalistica e che anche il terzo punto era aperto vale a dire – dal suo punto di vista – suscettibile di «ripensamento» da parte di Hilbert. È chiaramente il quarto punto quello cruciale, sul quale (a parte forse il caso isolato di von Neumann) i formalisti non erano disposti a cedere; solo due anni prima Hilbert aveva decisamente affermato: «No, se giustificare una procedura significa qualcosa di più che dimostrare la sua coerenza, ciò può solo significare determinare se la procedura stessa assolve i compiti per i quali è istituita.» Per concludere, almeno per ora, su quello che era l'aspetto in definitiva puramente «verbale» della polemica – prima cioè che intervenissero precisi *risultati* a determinarne meglio la portata – niente ci sembra più appropriato delle parole con cui Brouwer terminava l'articolo del '27:

«... Il formalismo non ha ricevuto che benefici dall'intuizionismo e ancora ne può attendere per il futuro. La scuola formalistica dovrebbe quindi accordare un certo riconoscimento all'intuizionismo, invece di polemizzare contro di esso con accenti beffardi... Inoltre la scuola formalistica dovrebbe riflettere sul fatto che nel contesto del formalismo finora non è stato propriamente garantito *niente* della matematica (poiché, dopo tutto, la dimostrazione matematica di coerenza per il sistema di assiomi manca, oggi come ieri) mentre l'intuizionismo... ha già edificato di bel nuovo molte delle teorie matematiche con certezza al di fuori di ogni dubbio. Se dunque la scuola formalistica... qualifica come modesto l'intuizionismo dovrebbe... non essere ad esso inferiore in questo riguardo.»

Ma veniamo ora brevemente all'articolo di Kolmogorov che, come si è più volte detto, consentì per la prima volta di stabilire un concreto collegamento fra l'aspetto logico dell'intuizionismo brouweriano e le altre concezioni (in particolare, ovviamente, con il formalismo) al di là delle enunciazioni verbali dei vari programmi. Kolmogorov prende le mosse dal sistema $\mathfrak{H}_1$ coerente e completo di logica proposizionale classica, presentato da Hilbert in *Die logischen Grundlagen der Mathematik* (*I fondamenti logici della matematica*, 1923), costituito dai sei assiomi

$$\left.\begin{array}{ll} 1) & p\rightarrow(q\rightarrow p) \\ 2) & (p\rightarrow(p\rightarrow q))\rightarrow(p\rightarrow q) \\ 3) & (p\rightarrow(q\rightarrow r))\rightarrow(q\rightarrow(p\rightarrow r)) \\ 4) & (q\rightarrow r)\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r)) \end{array}\right\} \text{assiomi dell'implicazione}$$

$$\left.\begin{array}{ll} 5) & p\rightarrow(\neg p\rightarrow q) \\ 6) & (p\rightarrow q)\rightarrow((\neg p\rightarrow q)\rightarrow q) \end{array}\right\} \text{assiomi della negazione}$$

con le regole di separazione e sostituzione.

Kolmogorov vuole stabilire «una logica generale del giudizio» e ritiene necessario disporre di principi (assiomi) che possano applicarsi a *ogni* giudizio (cioè ad ogni *proposizione*) e che quindi siano costituiti senza fare appello a nessun'altra proprietà dei giudizi stessi che non sia quella fondamentale stabilita già da Aristotele: il poter essere considerati dal punto di vista della verità o falsità. Ciò posto (dando un primo esempio di interpretazione costruttiva

dei connettivi), riconosce la validità per questa logica generale dei primi 4 assiomi $\mathfrak{H}_1$.[50] Le cose non stanno così per la negazione; questo connettivo (applicato a un giudizio) può essere interpretato, dice Kolmogorov, in due modi: 1) come proibizione di considerare quel giudizio come vero; 2) pensando il giudizio come l'attribuzione di un predicato a un soggetto, come l'affermazione che il predicato è incompatibile col soggetto. Nell'ambito della logica generale che Kolmogorov vuol costruire solo la prima interpretazione è corretta;[51] ma l'«usuale tradizione in logica è stata quella di passare dalla prima interpretazione alla seconda, riguardata come più primitiva. Nell'applicazione ai giudizi matematici ciò risulta essere impossibile». Ne discende che nessuno dei due assiomi della negazione del sistema $\mathfrak{H}_1$ può essere mantenuto (si noti che il secondo esprime proprio la legge del terzo escluso). Egli allora propone un sistema $\mathfrak{B}$ (da Brouwer) costituito dai primi 4 assiomi di Hilbert con l'aggiunta dell'unico assioma

$$(7) \qquad (p \to q) \to ((p \to \neg q) \to \neg p)$$

che è adeguato appunto alla prima interpretazione sopra data della negazione come assurdità.[52] Ora egli indica con $\mathfrak{H}$ (da Hilbert) il sistema costituito dagli assiomi 1, 2, 3, 4, 7, 8; si ha così il vantaggio che il sistema $\mathfrak{H}$ viene ottenuto «dal sistema $\mathfrak{B}$ della logica generale dei giudizi con la sola aggiunta dell'assioma della doppia negazione; *ciò facilita considerevolmente ulteriori ricerche*». Il sistema $\mathfrak{B}$ di Kolmogorov non coincide con quello che darà Heyting qualche

[50] Essi costituiscono un sistema positivo implicazionale e Kolmogorov si chiede nel suo articolo (lasciando aperta la questione) se tale sistema sia completo, ossia se tutte le implicazioni tautologiche positive (cioè senza negazione) possano essere ottenute a partire da quei quattro assiomi. Oggi sappiamo che la risposta alla questione è negativa, in quanto per ottenere un sistema completo occorre aggiungere ai quattro assiomi un ulteriore principio, detto legge di Peirce: $((p \to q) \to p) \to p$.

[51] Essa, come dice Kolmogorov, fu introdotta per la prima volta da Brouwer nel 1923. Questi definisce «la negazione come assurdità» e si fonda sulla seconda interpretazione «poiché per derivare un giudizio negativo riducendolo a contraddizione dobbiamo già disporre di qualche giudizio negativo»; ma è nello stesso tempo più ampia di questa interpretazione.

[52] Kolmogorov fa anche vedere come non sia intuizionisticamente accettabile la legge della doppia negazione (8) $\neg\neg p \to p$ (che di fatto è equivalente al terzo escluso) mentre sia derivabile nel sistema $\mathfrak{B}$ l'implicazione opposta $p \to \neg\neg p$.

anno dopo, ma con quello presentato nel 1936 da Ingebrigt Johansson e oggi noto come *sistema minimale*, ma il fatto importante è che mostra quali principi logici ci facciano passare da un sistema costruttivo come $\mathfrak{B}$ ad uno classico come $\mathfrak{H}_1$. Kolmogorov non si preoccupa di dare una analoga versione completa della logica dei predicati intuizionista, malgrado dalle considerazioni che svolge in proposito si possa facilmente ottenere quello che sarà il sistema di Heyting salvo un assioma (e precisamente l'assioma 5 di Hilbert che Heyting invece accetterà).

Va comunque osservato che l'idea veramente centrale ed estremamente importante che muoveva il lavoro di Kolmogorov consisteva nel prospettare (anche se in modo non sempre limpido ed esplicito) la possibilità di una traduzione della matematica classica nella matematica intuizionista. Ancora oggi, osserva Hao Wang, «ciò non è stato stabilito con certezza per quanto riguarda l'Analisi classica e la teoria degli insiemi. D'altra parte, non sembra irragionevole affermare che Kolmogorov prevedeva che il sistema della teoria dei numeri classica è traducibile nella corrispondente teoria intuizionistica e quindi è intuizionisticamente coerente». Ma riprenderemo questo argomento nel prossimo capitolo mostrando come le ricerche di Kolmogorov convergano con quelle di Heyting nell'elaborazione di una vera e propria semantica costruttiva per la logica intuizionista.

3.6 *Conclusione sugli anni venti*

Tirando le fila del discorso svolto in questo capitolo, si può affermare che la situazione della ricerca logica e sui fondamenti attorno agli anni venti va perdendo quel carattere di «fluidità» che l'aveva ancora caratterizzata nei primi anni del decennio e acquisisce invece alcuni punti fermi: la sistemazione della teoria degli insiemi divenuta ormai un riferimento centrale della ricerca, sia pure solo come sistema fondazionale (restando peraltro alquanto «sterile», fino a questo momento, come puro sistema matematico); il deciso imporsi di due dei tre programmi di fondazione della matematica, il formalismo e l'intuizionismo, che pur nelle violente polemiche presentano tuttavia maggiori e stretti collegamenti di quanto i sostenitori stessi dei due indirizzi non siano forse disposti

ad ammettere: il primo quasi in attesa della definitiva consacrazione, il secondo come pressante, coerente e «ragionevole» richiamo a una matematica «reale» senza compromessi.

In questo sfondo acquista quasi un valore simbolico il simposio tenuto a Königsberg nel 1930 sui fondamenti della matematica, cui parteciparono rappresentanti autorevoli dei tre indirizzi presentando tre memorie pubblicate nel 1931 sulla rivista *Erkenntnis:* Rudolf Carnap con *Die logizistische Grundlegung der Mathematik* (*La fondazione logicista della matematica*), Arend Heyting, con *Die intuizionistiche Grundlegung der Mathematik* (*La fondazione intuizionista della matematica*) e John von Neumann con *Die formalistische Grundlegung der Mathematik* (*La fondazione formalista della matematica*).

È subito chiaro che il compito più difficile spetta a Carnap: deve difendersi da troppe obiezioni e giustificare profonde difficoltà intrinseche all'approccio logicista e deve nel contempo stabilire più stretti rapporti con le altre due concezioni da cui sostanzialmente è stata emarginata quella da lui rappresentata. È proprio per questo che Carnap insiste sul «costruttivismo» del metodo logicista, il quale non *postula* ma *costruisce* i numeri reali e tale tendenza costruttivistica – che costituisce a suo parere un fondamentale punto di contatto con l'intuizionismo[53] – non viene inficiata dalla presenza delle definizioni impredicative, almeno se si accetta l'analisi che di queste fa Carnap. Più stretto ancora se possibile – sempre a parere di Carnap – il collegamento con i formalisti che, oltre a esplicarsi a livello metodologico, diventa vera coincidenza *all'interno* dei sistemi concreti: qui anche i logicisti sono formalisti. Si noti che Heyting non fa il minimo accenno né alla posizione formalista, né tanto meno, alla logicista: si limita a esporre succintamente le posizioni intuizioniste specie ora che egli stesso ne ha dato una formulazione «formale»; propone un'interpretazione dei connettivi logici basata sulla concezione (che egli fa risalire alla fenomenologia) di una proposizione come «intenzione» (di una costruzione) che viene linguisticamente espressa dalla proposizione stessa e con molta semplicità e altrettanta sicurezza conclude: «avremo raggiunto il nostro scopo se avremo mostrato che l'intuizionismo non contiene assunzioni arbitrarie. Ancor meno esso

[53] Carnap ammette tuttavia anche l'esistenza di notevoli divergenze tra i due indirizzi: i logicisti accettano, a differenza degli intuizionisti, logiche del secondo ordine e con forti principi esistenziali.

contiene proposizioni artificiali come quelle usate per evitare i paradossi logici. *Piuttosto, una volta adottata la sua attitudine di base, l'intuizionismo è l'unica via possibile per costruire la matematica*» [corsivo nostro].

Von Neumann viceversa (che potremmo definire il «più intuizionista» dei formalisti) mette in luce una specie di collegamento operativo tra i risultati dei vari indirizzi ascrivendo tuttavia in particolare alle recenti indagini dei formalisti il merito della trasformazione dei problemi relativi ai fondamenti della matematica da vaghe questioni di carattere filosofico-epistemologico in problemi precisi: «Come risultato di tre importanti progressi nella logica matematica (precisamente: la sottile formulazione di Brouwer dei difetti della matematica classica; la minuziosa ed esatta descrizione di Russell dei suoi metodi (tanto i buoni quanto i cattivi); e i contributi di Hilbert alla ricerca matematico-combinatoria di questi metodi e delle loro relazioni) sono sempre più questioni matematiche non ambigue, non questioni di gusto, che devono essere investigate nei fondamenti della matematica.» Dà quindi un'esemplare e stringata caratterizzazione del metodo hilbertiano e della sua idea di fondo: «... anche se gli enunciati della matematica classica dovessero risultare falsi quanto a contenuto, cionondimeno la matematica classica comporta una procedura internamente chiusa che opera secondo regole fisse note a tutti i matematici e che consiste fondamentalmente nel costruire successivamente certe combinazioni di simboli primitivi che sono considerate "corrette" o "dimostrate". Questa costruzione inoltre è "finitaria e direttamente costruttiva"». Pone quindi l'equazione finitismo = intuizionismo e ne conclude che «sebbene il contenuto di una proposizione matematica classica non possa essere sempre (ossia in generale) verificato finitamente, il modo formale col quale noi arriviamo alla proposizione lo può. Di conseguenza, se vogliamo dimostrare la validità della matematica classica – il che in linea di principio è possibile solo riducendola al sistema finitistico valido *a priori* (ossia al sistema di Brouwer) – allora dobbiamo investigare non enunciati, ma metodi di dimostrazione. Dobbiamo riguardare la matematica classica come un gioco combinatorio giocato con i simboli primitivi e dobbiamo determinare in modo combinatorio a quali combinazioni di simboli primitivi conducono i metodi di costruzione o "dimostrazioni"». Pur concludendo

per così dire con una professione di fede formalista è chiaro quindi che von Neumann assegna un posto di primo piano alle indagini e alla stessa concezione intuizionista.

Vogliamo solo ricordare brevemente, per finire, le proposte di Carnap per superare le difficoltà del logicismo. Allo scopo Carnap distingue nel programma logicista una parte relativa ai *concetti* della matematica (che possono essere ottenuti – secondo la tesi logicista – da concetti logici mediante definizioni esplicite, cioè costruzioni di tipo insiemistico) e una parte relativa ai *teoremi* della matematica (che possono essere derivati da assiomi logici mediante deduzioni puramente logiche). Per quanto riguarda il primo punto (essenzialmente la definizione dei concetti di numero naturale, reale ecc.), precisati i concetti logici di fondo, Carnap riconosce che la definizione di numero reale presenta difficoltà che a suo parere «né il logicismo, né l'intuizionismo, né il formalismo hanno ancora superato». In particolare un grave problema in questo senso è quello delle definizioni impredicative. Per quanto riguarda il secondo punto si prospettano almeno tre grosse difficoltà.

La prima è la necessità delle assunzioni esistenziali (infinito e scelta) richieste per la dimostrazione di alcuni teoremi; e la sua gravità è data dal fatto che «la logica tratta solo con entità possibili, e non può fare asserzioni sul fatto se qualcosa esiste o non esiste». Carnap comunque ritiene di poter aggirare questa difficoltà ricorrendo al discorso ipotetico (ossia condizionale in quanto opposto a categorico): *se* esistono certe strutture, *allora* si può concludere logicamente all'esistenza di certe altre.

La seconda, maggiore difficoltà è ovviamente rappresentata dall'assunzione dell'assioma di riducibilità, che viene da Carnap riportata – via discorso sulla ramificazione – alla terza questione problematica, rappresentata dalle definizioni impredicative.[54]

A questo proposito Carnap non accetta la soluzione di Ramsey che trova la sua base in un realismo concettuale, o, come dice Car-

[54] Carnap illustra le difficoltà che la ramificazione comporta per la teoria dei numeri reali dal momento che non si può avere né un'ammissibile definizione né un'ammissibile proposizione che si riferisca a tutti i numeri reali senza qualificazione di ordine, e così enuncia quello che per lui è il più difficile problema dei fondamenti: «come possiamo sperare di sviluppare la logica se da una parte dobbiamo evitare il pericolo di definizioni prive di significato o impredicative e dall'altra dobbiamo ricostruire in modo soddisfacente la teoria dei numeri reali [Analisi]?».

nap, ipotizza un «mondo platonico ove le idee sussistono di per sé... indipendentemente da *se* e *come* le pensiamo». E contrappone alla «matematica teologica» di Ramsey la «matematica antropologica» degli intuizionisti, con i quali in particolare concorda sul fatto che «la finitezza di ogni operazione, dimostrazione o definizione logico-matematica non è richiesta a causa di alcuni fatti empirici accidentali nell'uomo, ma per la natura stessa del soggetto». Si tratta allora di ottenere i vantaggi della soluzione di Ramsey evitandone tuttavia la portata assolutista. Orbene, «se rigettiamo la credenza che è necessario passare attraverso tutti i casi individuali e piuttosto rendiamo chiaro a noi stessi che la verifica completa di un enunciato su una proprietà arbitraria non significa nulla di più che la sua validità logica (più precisamente tautologica) per nulla arbitraria, giungeremo alla conclusione che le definizioni impredicative sono logicamente ammissibili». Il fatto che si sappia o meno decidere se una proprietà definita impredicativamente sia valida o no per un caso particolare può dipendere da molte cose, ma mai dalla impredicatività; e se questa sua soluzione è valida, il logicismo è riuscito a liberarsi dalla stretta di Scilla (assioma di riducibilità) e Cariddi (vari ordini di numeri reali).

Ma, si può dire, il gioco era ormai fatto. I riferimenti erano ormai intuizionismo e formalismo: il decennio successivo decreterà addirittura anche la «fine» del formalismo, almeno inteso in senso hilbertiano; e la grande avventura logicista, in effetti, viveva già soltanto della sua storia gloriosa. Ben presto la situazione avrebbe assunto caratteri nuovi e inaspettati. È infatti al convegno di Königsberg che Gödel presentò per la prima volta i suoi risultati sulla indecidibilità dell'aritmetica.

B. Russell, *I principi della matematica*, Longanesi, Milano 1951 [edizione originale 1903].

Id., *Introduzione alla filosofia della matematica*, Longanesi, Milano 1947 [edizione originale 1919].

Id. e A.N. Whitehead, *Principia Mathematica*, Cambridge University Press, Cambridge, 3 volumi, 1910, 1927².

Id., *Introduzione ai Principia Mathematica*, a cura di Paolo Parrini, La Nuova Italia 1967.

B. Russell, *Collected Papers*, in corso di pubblicazione presso Allen and Unwin, Londra.

F.P. Ramsey, *I fondamenti della matematica e altri scritti di logica*, Feltrinelli, Milano 1964 [edizione originale 1931].

Id., *Notes on Philosophy, Probability and Mathematics*, a cura di Maria Carla Galavotti, Bibliopolis, Napoli 1991.

A. Fraenkel e Y. Bar-Hillel, *Foundations of Set Theory*, North Holland, Amsterdam 1973.

A.A. Fraenkel, *Abstract Set Theory*, North Holland, Amsterdam 1968.

P. Bernays, *Axiomatic Set Theory*, North Holland, Amsterdam 1958.

K. Gödel, *The consistency of the axiom of choice and of the generalized continuum Hypothesis with axioms of Set Theory*, Princeton University Press, Princeton 1966 [edizione originale 1940].

G.H. Moore, *Zermelo's Axiom of Choice: its Origins, Development an Influence*, Springer Verlag, Berlino 1980.

S.H. Moore, *The Emergence of First-Order Logic* in W. Aspray e P. Kitcher (a cura di), *History and Philosophy of Modern Mathematics*, University of Minnesota Press, Minneapolis 1888, pp. 95-135.

F. Hausdorff, *Set Theory*, Chelsea, New York 1962 [edizione originale 1937].

J. von Neumann, *Collected Works*, volume I, a cura di A.H. Taub, Pergamon, Oxford 1961.

L.E.J. Brouwer, *Collected Works*, 2 volumi, a cura di Arend Heyting, North Holland, Amsterdam 1975.

Id., *Lezioni sull'intuizionismo*, Boringhieri, Torino 1983 [edizione originale 1981].

A. Heyting, *Intuitionism. An Introduction*, North Holland, Amsterdam 1971.

W.P. van Stigt, *Brouwer's Intuitionism*, North Holland, Amsterdam 1990.

E. Beth, *I fondamenti logici della matematica*, Feltrinelli, Milano 1963 [edizione originale 1955].

H. Weyl, *Il continuo*, Bibliopolis, Napoli 1977 [edizione originale 1917].

Id., *Filosofia della matematica e delle scienze naturali*, Boringhieri, Torino 1967 [edizione originale 1949].

AA.VV., *Dalla logica alla metalogica*, a cura di Ettore Casari, Sansoni, Firenze 1979.

J. Łukasiewicz, *Selected Works*, North Holland, Amsterdam 1970.

AA.VV., *Polish logic*, a cura di S. McCall, Clarendon Press, Oxford 1967.

L. Post, *The Two-Valued Iterative Systems of Mathematical Logic*, Princeton University Press, Princeton 1941.

C.I. Lewis, op. cit.

C.I. Lewis e C.H. Langford, *Symbolic Logic*, Dover Publications, Inc., New York 1959 [edizione originale 1932].

D. Hilbert e W. Ackermann, *Principles of mathematical Logic*, Chelsea, New York 1959 [edizione originale 1928].

D. Hilbert e P. Bernays, *Grundlagen der Mathematik*, 2 volumi, Springer Verlag, Berlino 1968 [edizione originale 1934].

AA.VV., *Philosophy of Mathematics*, a cura di P. Benacerraf e H. Putnam, Prentice Hall, New Jersey 1964.

AA.VV., *La filosofia della matematica*, a cura di Carlo Cellucci, Laterza, Bari 1967.

R. Carnap, *Fondamenti di logica e matematica*, a cura di Giulio Preti, Paravia, Milano 1956 [edizione originale 1939].

Th. Skolem, *Selected Works in Logic*, a cura di Jens Erik Fenstad, Universitetsforlaget, Oslo 1970.

G. Lolli, *Le ragioni fisiche e le dimostrazioni matematiche*, Il mulino, Bologna 1985.

E. Casari, *Questioni di filosofia della matematica*, Feltrinelli, Milano 1964.

CAPITOLO QUINTO
LA SVOLTA DEGLI ANNI TRENTA

1. INTRODUZIONE

Gli anni trenta si aprono con un risultato, il teorema di completezza semantica di Gödel (1930) per la logica del primo ordine, che sembrava confortare anche da un punto di vista generale la convinzione che di tutti i programmi fondazionali, quello hilbertiano rappresentasse la via della ragionevolezza, coniugando adeguatezza alla pratica matematica con correttezza intuitiva. Questo, anche se esso risultava strettamente legato ad un altro teorema dimostrato da Löwenheim nel 1915 e variamente elaborato da Skolem nel '20 e nel '22 (oggi conosciuto come teorema di Löwenheim-Skolem) il cui carattere limitativo nei confronti della logica del primo ordine abbiamo avuto modo di sottolineare nel capitolo precedente. Ma la sorte decisiva del programma hilbertiano sarebbe stata segnata solo un anno più tardi da un secondo, importantissimo lavoro di Gödel, che conteneva quello che viene detto il «teorema di Gödel» per antonomasia e che decretava l'impossibilità di principio di una dimostrazione finitista della coerenza della teoria dei numeri. È su questo sfondo che si situano gli sviluppi più decisivi della logica di questi anni, in particolare la nascita della teoria della computazione e della decidibilità che nei risultati di Gödel trovò l'iniziale spinta propulsiva, la già varie volte citata «formalizzazione» della logica e della matematica intuizioniste ad opera di Heyting e infine la sistemazione rigorosa della semantica dei linguaggi (dei sistemi) formali ad opera di Tarski.

Per inquadrare adeguatamente la situazione ed afferrare il senso del teorema di completezza di Gödel, è opportuno fare un passo indietro nella nostra storia e richiamare brevemente lo sviluppo di quell'approccio algebrico alla logica che era culminato con Schröder e come sappiamo si caratterizzava come approccio *semantico*

alla logica (in contrasto con l'altro approccio – di tipo logicista e formalista – che pur con sfumature e concezioni diverse privilegia una via sostanzialmente sintattica).

L'approccio di Schröder, si ricorderà, assume – se pure a livello intuitivo e non rigorosamente teorizzato – una «gerarchia» dei linguaggi, che sfruttando estensioni infinitarie (somme e prodotti logici infiniti) e altri tipi di sviluppo di ispirazione algebrica poneva in luce la grande flessibilità ed espressività di quello che noi abbiamo chiamato «linguaggio predicativo del primo ordine» (con o senza identità). Le proprietà generali di questi linguaggi, o delle teorie tramite essi descritte, venivano indagate non in termini delle nozioni sintattiche di «assioma», «dimostrazione», «regola di inferenza», «noncontraddittorietà» e così via, bensì in termini di nozioni semantiche *intuitive* quali quelle di «dominio di interpretazione», «soddisfacibilità», «validità», «verità», ecc. In particolare la stessa nozione di contraddittorietà non veniva assunta nel suo significato sintattico, bensì in quello semantico di «non soddisfacibilità» (e analogamente la nozione di noncontraddittorietà veniva invece intesa come soddisfacibilità).

Questi termini semantici non avevano ovviamente nulla di misterioso e venivano intesi nel loro immediato significato intuitivo: se abbiamo una teoria espressa nel linguaggio del primo ordine, questa può essere *interpretata* in domini diversi, vale a dire si può assumere un insieme di oggetti noti di qualche altra teoria, per lo più i numeri naturali, come universo del discorso della teoria e collegare (*interpretare*, appunto) i segni linguistici con (su) elementi del dominio o con (su) relazioni o operazioni fra elementi del dominio e così via. Così facendo, le formule del linguaggio, in particolare le formule chiuse (ossia senza variabili libere) divengono proposizioni relative al dominio stesso e come tali possono risultare vere o false. Intuitivamente, una *formula* è vera in un certo dominio se tramite un'opportuna interpretazione del linguaggio (in cui la formula è espressa) su quel dominio essa si trasforma appunto in una *proposizione* vera (in quel dominio) mentre sarà in generale soddisfacibile se esiste un dominio che via interpretazione opportuna la rende vera; non soddisfacibile nel caso contrario. Infine sarà *valida* se comunque si scelgano un dominio e un'interpretazione del linguaggio su di esso la formula in questione risulterà vera. Analogamente, una teoria (ossia un insieme di formule) sarà

soddisfacibile (noncontraddittoria) se esiste un dominio che soddisfa i suoi assiomi (che rende vere tutte le formule che costituiscono gli assiomi della teoria), non soddisfacibile (contraddittoria) nel caso contrario. Nel caso in cui la teoria sia soddisfacibile su un dato dominio si dice che essa ammette un *modello*, costituito da quel dominio con le relazioni e operazioni opportunamente definite su di esso.

Il problema non era che tutta questa concettualizzazione semantica veniva usata in senso largamente intuitivo: i logici e i matematici «si intendono» quando usano termini semantici di quel tipo, che d'altra parte sono preferibili ai concetti sintattici puramente formali perché sono contenutisticamente più rilevanti. Ciò che sollevava difficoltà a livello epistemologico era il ricorso, in essi implicito, a concetti insiemistici come ad esempio quello di dominio, di relazione, ecc. Che significava *dare* una interpretazione? Caso per caso era chiaro, ma non c'era pieno accordo su una caratterizzazione generale di questo concetto. Nessuna meraviglia allora che non si ponesse per questi ricercatori la questione del rapporto fra i due approcci possibili allo studio delle teorie, ossia che non ci si ponesse neppure implicitamente la questione di un'eventuale coincidenza fra formule (semanticamente) «valide» e formule (sintatticamente) «dimostrabili», in quanto l'interpretabilità non era vista come un concetto determinato in modo adeguato. E si noti che, in generale, un problema (o una dimostrazione) di completezza (semantica) consiste nel chiedersi se (nel dimostrare che) ogni formula valida è (formalmente) dimostrabile in un dato sistema formale.

In questo ordine di idee Leopold Löwenheim (1878-1943) aveva dimostrato nel 1915 in *Über Moglichkeiten im Relativkalkül* (*Sulle possibilità di calcolo dei relativi*; si noti la terminologia peirciana) un teorema che possiamo enunciare come segue: se *una* formula del primo ordine è soddisfacibile, allora tale formula è soddisfacibile in un dominio infinito numerabile. Vale a dire, se mai si può trovare un dominio D in cui una formula $\mathscr{A}$ («scritta» in un linguaggio) del primo ordine è soddisfacibile, allora esiste un dominio D' che contiene un'infinità numerabile di elementi (e che può quindi essere identificato una volta per tutte, ad esempio, con l'insieme N dei numeri naturali) in cui $\mathscr{A}$ è soddisfacibile («interpretando» opportunamente le costanti predicative contenute in $\mathscr{A}$).

Nel 1920, 1922 (e in altre occasioni) Skolem diede ulteriori dimostrazioni e generalizzazioni del teorema di Löwenheim sensibilmente diverse. Nella memoria del 1920 *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen* (*Ricerche logico-combinatorie sulla soddisfacibilità o dimostrabilità delle proposizioni matematiche e un teorema sugli insiemi densi*) Skolem compiva un passo decisivo e sottolineava come la dimostrazione di Löwenheim assumesse forma semplificata una volta accettato che i domini, le interpretazioni, ecc., sono oggetti insiemistici e come tali vanno trattati utilizzando tutti gli strumenti della teoria degli insiemi. Riconosciuto però che il teorema di Löwenheim era «fondato sulla teoria degli insiemi», Skolem non contrapponeva però, malgrado il titolo, alla nozione insiemistica di soddisfacibilità una nozione di dimostrabilità con precisi caratteri sintattici, non riuscendo a vedere nel rapporto fra concetti semantici e sintattici un problema da indagare formalmente ponendosi la domanda della completezza degli schemi inferenziali.

L'enfasi di Skolem era sui metodi per la costruzione di modelli e la tecnica usata in questo come in successivi lavori si basava sull'idea che, mentre nel caso di formule che contengono solo connettivi è immediato costruire valutazioni modello ricorrendo alle regole di computo date dalle tavole di verità, il problema con i linguaggi del primo ordine è la presenza dei quantificatori che, facendo riferimento a domini d'interpretazione arbitrari, introducono un elemento infinitario non immediatamente assorbibile entro un contesto combinatorio. Per fare questo era quindi necessario per prima cosa separare i quantificatori dai connettivi, poi eliminare i quantificatori esemplificando gli esistenziali con sempre nuove costanti individuali e ricondursi così al caso proposizionale.

Skolem partiva dal lavoro di Löwenheim che aveva affrontato la questione in una maniera che – esprimendoci in modo divenuto usuale – possiamo schematizzare nel modo seguente. Per prima cosa associamo ad ogni formula $\mathscr{A}$ una $\mathscr{A}'$ equivalente ad essa in *forma prenessa*. In generale diciamo *prenessa* ogni formula $\mathscr{A}$ del tipo

$$Q_1 x_{11} \ldots x_{m1} \ldots Q_k x_{1k} \ldots x_{nk} \mathscr{B}$$

dove $\mathscr{B}$ non contiene quantificatori (ed è chiamata la *matrice* di $\mathscr{A}$, seguendo una terminologia già introdotta da Russell e Whitehead) e i vari Q_j, $j \leqslant k$, sono quantificatori $\forall$ o $\exists$. Ogni successione $Q_j x_{11} \ldots x_{mj}$ costituisce un blocco in cui uno stesso tipo di quantificatore vincola la successione di variabili $x_{11} \ldots x_{mj}$. La successione di tutti i blocchi prima di $\mathscr{B}$ viene detta *prefisso* di $\mathscr{A}$ e costituisce una misura significativa della complessità dei riferimenti che la formula fa alla totalità del dominio di interpretazione. Non è infatti il numero delle successive quantificazioni di uguale tipo (tutte universali o tutte esistenziali) che conta, ma *l'alternarsi* di universali ed esistenziali che dà la complessità. Ad esempio la formula

$$\forall x \exists y \mathscr{B}(x, y)$$

è sicuramente più complessa delle formule

$$\forall x \forall y \mathscr{B}(x, y) \text{ oppure } \exists x \exists y \mathscr{B}(x, y)$$

in quanto per essere verificata richiede che per ogni a del dominio D di una possibile interpretazione esista un $b \in D$ per cui $\mathscr{B}(a, b)$. Se vogliamo costruire un modello del genere dobbiamo scegliere b *in funzione* di a. Questo non succede per le altre due formule in cui le scelte per x e y possono essere indipendenti e quindi *simultanee*, mentre nel primo caso occorre procedere un passo dopo l'altro. L'esperienza tanto nello studio dell'*Entscheidungsproblem*, quanto nell'analisi dei modelli delle formule aveva posto in luce l'importanza di una simile distinzione, come pure il fatto che per ogni formula $\mathscr{A}$ ne esiste una prenessa equivalente (con le stesse variabili libere) come si può provare facilmente sfruttando la distributività di connettivi e quantificatori (teorema di forma prenessa).

Dopo questo primo passo, Löwenheim passa ad eliminare i quantificatori esistenziali esemplificandoli. Ad esempio, partendo dalla formula prenessa

$$\mathscr{A} \qquad \forall x \exists y \forall z \mathscr{B}(x, y, z)$$

introduceva un simbolo di funzione f unario e considerava la formula

$$\mathscr{A}_S \qquad \forall x \forall z \mathscr{B}(x, f(x), z).$$

È chiaro che se $\mathscr{A}_S$ (oggi detta *skolemiana* di $\mathscr{A}$) è soddisfacibile lo è anche $\mathscr{A}$: dato un qualsiasi a nel dominio D, se $\mathscr{A}_S$ è vera nell'interpretazione, $f(a)$ sarà tale che anche $\mathscr{B}(a, f(a), b)$ sarà vera

per ogni $b \in D$, così che risulta verificata la $\mathscr{A}$. La funzione denotata da f rende esplicita la dipendenza di y da x che avevamo sottolineato sopra. In questo modo ad ogni formula prenessa $\mathscr{A}$ possiamo associare una formula prenessa $\mathscr{A}_S$ che ha un prefisso composto da soli quantificatori universali. La formula contiene nuovi simboli per funzioni che esemplificano gli esistenziali, ciascuno di tanti argomenti quante sono le variabili quantificate precedentemente con $\forall$, ed ha struttura più semplice. Il problema è ora quello di provare l'inverso, cioè che se $\mathscr{A}$ è soddisfacibile lo sarà anche $\mathscr{A}_S$. Questo non si può fare – come ribadito da Skolem – senza appellarsi in qualche modo all'assioma di scelta che appunto ci permette di affermare che per ogni relazione binaria R se vale che $\forall x \exists y R(x,y)$ esiste una funzione f per cui $\forall x R(x, f(x))$. Questa formulazione dell'assioma risale a Hilbert e riflette appunto l'idea che i quantificatori si possono rimpiazzare con un operatore come l'ε-operatore che abbiamo incontrato nel capitolo IV, e che coinvolge un'operazione di scelta dei testimoni. Di fatto Löwenheim usava l'assioma senza farvi esplicito riferimento, atteggiamento questo comune a tutta la scuola di Schröder che, usando una terminologia modellata su quella algebrica, parlava di una *legge logica generale* (una legge distributiva) per cui

$$\forall x \exists y \mathscr{B}(x,y) \longleftrightarrow \exists f \forall x \mathscr{B}(x, f(x))$$

il che è un modo appunto di formulare la scelta.

Curiosamente, le funzioni così introdotte sono note oggi come *funzioni di Skolem*, anche se di fatto solo in lavori successivi a quello del 1920 (come quello del 1928 dal titolo *Über die mathematische Logik* [*Sulla logica matematica*]) Skolem le utilizzò effettivamente. La via seguita nel lavoro del '20 si basava sull'idea di associare ad ogni formula come la $\mathscr{A}$ di sopra, un'altra formula $\mathscr{A}^*$ prenessa, la cui complessità è vincolata nel senso che il prefisso sarà del tipo $\forall\ \exists$ (o come oggi si scrive del tipo Π^0_2) senza introdurre costanti funzionali. Questa formula $\mathscr{A}^*$ viene oggi chiamata *forma normale di Skolem per la soddisfacibilità* e si ha che

$$\mathscr{A} \text{ è soddisfacibile se e solo se } \mathscr{A}^* \text{ lo è.}$$

Skolem provava allora che ogni formula $\mathscr{A}$ ha una forma normale e il metodo è quello di introdurre al posto di simboli funzionali simboli per relazioni; esemplificando per la $\mathscr{A}$ di sopra, Skolem mostra che $\mathscr{A}$ è soddisfacibile se e solo se lo è

$$\forall x \exists w R(x, w) \wedge \forall x \forall u (R(x, u) \longleftrightarrow \forall y \mathscr{B}(x, u, y))$$

dove R è la nuova costante introdotta per eliminare $\exists w$. Si verifica quindi che $\mathscr{A}$ è soddisfacibile se e solo se lo è

$$\forall x \forall u \forall y \exists w \exists z (R(x, w) \wedge ((R(x, u) \wedge \mathscr{B}(x, u, y)) \vee (\neg R(x, u) \wedge \mathscr{B}(x, u, z)))).$$

Diversamente da Löwenheim, Skolem, come si è detto, era consapevole che per mostrare questa equivalenza è necessario l'assioma di scelta e la sua critica a Löwenheim riguardava solo l'eccessiva complicazione della dimostrazione. Di fatto però, una volta passati alla forma normale, la tecnica seguita da Skolem per costruire un modello di $\mathscr{A}^*$ sarà la stessa, se non nell'articolo del '20 in quello del 1929 *Über einige Grundlagenfragen der Mathematik* (*Su alcune questioni circa i fondamenti della matematica*). Il problema comune rimaneva quello di costruire un'interpretazione il cui dominio fosse dato (facendo riferimento alla formula $\mathscr{A}_S$ di sopra) dai termini del linguaggio arricchito dai nuovi simboli funzionali; si trattava di definire una interpretazione di tutte le formule atomiche del tipo $P(t_1 \ldots t_n)$ in modo che risultassero veri tutti gli esempi della matrice di $\mathscr{A}_S$ ottenuti sostituendo quei termini alle variabili. Poiché il dominio è dato proprio da questi termini, $\mathscr{A}_S$ risulterà vera nell'interpretazione data e per ottenere quest'ultima sarà sufficiente muoversi a livello combinatorio *proposizionale* poiché la matrice non contiene quantificatori.

È questo un aspetto della costruzione dei modelli che ritroveremo nella dimostrazione di completezza di Gödel, anch'essa basata su un dominio costituito da oggetti linguistici. L'idea è di procedere livello per livello al crescere della complessità dei termini: si parte da un primo livello L_1 con solo le costanti individuali, si passa ad un secondo livello L_2 in cui si applicano le costanti funzionali ai termini di L_1 e così via. Ad ogni livello avremo un numero finito di termini chiusi e finito sarà quindi anche il numero degli esempi della matrice di $\mathscr{A}$ costruiti con questi termini. Se chiamiamo *soluzione di livello n* ogni valutazione che soddisfa tutti gli esempi di livello n, si può ora provare che per ogni n esiste un numero finito di soluzioni di livello n e che, per ogni n, almeno una soluzione di livello n si prolunga in una di livello $n + 1$. Possiamo così disporre le soluzioni in un albero infinito con diramazioni finite

(quello che più avanti, parlando di Brouwer, chiameremo *ventaglio*) ed un lemma dimostrato da König nel 1927 garantisce che tale albero avrà almeno un ramo infinito, vale a dire una valutazione di *tutti* gli esempi di ogni livello che ci darà un modello di $\mathscr{A}_S$. Questo, beninteso, nell'ipotesi che $\mathscr{A}_S$ sia soddisfacibile. Il teorema di Löwenheim scende allora considerando che per definizione il dominio è numerabile.

Anche se, come abbiamo ricordato, Skolem utilizzò metodi analoghi nel 1928 per provare l'adeguatezza di una sua procedura di refutazione basata sulla costruzione di contromodelli, egli non notò il passo cruciale della globalizzazione via lemma di König e questo perché qui Skolem poneva l'accento sulla *dimostrabilità* senza però basarsi su un calcolo con regole sintattiche ben precise, ma confondendo sistematicamente procedure inferenziali e metodi per la costruzione di contromodelli.

Ben altro atteggiamento si ricorderà era stato quello del 1920 quando, ponendosi su un piano dichiaratamente semantico, Skolem aveva dimostrato una generalizzazione del teorema di Löwenheim per cui ogni insieme numerabile di formule se soddisfacibile ha un modello numerabile. Qui Skolem utilizzava francamente metodi insiemistici mostrando come, dato un insieme T di formule, fosse possibile costruire un modello di T^* (l'insieme delle forme normali degli elementi di T) generalizzando il metodo delle catene di Dedekind: assumendo la scelta si otteneva così il desiderato modello di T. Più tardi, procedendo per via *sintattica*, Herbrand avrebbe dimostrato come non ricorrere all'assioma.

A prescindere dai grandi sviluppi che in seguito avrebbero avuto i teoremi di tipo Löwenheim-Skolem[1], la grande importanza del

[1] Essi infatti sono stati al centro di tutta una serie di indagini dedicate dagli anni venti in poi al problema della categoricità delle teorie formalizzate, cioè della possibilità di ottenere teorie con un modello solo, unico a meno di isomorfismi (e abbiamo visto che lo stesso Skolem lo utilizzava in questo senso). Il teorema di Skolem, infatti, è uno dei teoremi fondamentali della logica matematica e ha ottenuto numerose generalizzazioni che si possono sostanzialmente dividere in due gruppi: i) teoremi di Löwenheim-Skolem «superiori» (*upward* nella terminologia inglese) e ii) teoremi di Löwenheim-Skolem «inferiori» (*downward* nella terminologia inglese). I primi stabiliscono l'esistenza, per ogni modello infinito $\mathfrak{M}$ di una teoria $\mathfrak{T}$ (espressa entro un dato linguaggio) di sue estensioni $\mathfrak{M}'$, ancora modelli di $\mathfrak{T}$ e aventi cardinalità $\mathfrak{t}$ per ogni $\mathfrak{t}$ maggiore della cardinalità $\mathfrak{m}$ di $\mathfrak{M}$. I secondi stabiliscono l'esistenza di sottostrutture del modello $\mathfrak{M}$ che siano ancora modelli della teoria di partenza $\mathfrak{T}$ e che abbiano cardinalità $\mathfrak{t}$ infinita minore o uguale alla

risultato sul piano filosofico stava nelle conseguenze che Skolem ne trasse: il teorema dimostrava la radicale impossibilità di caratterizzare per via assiomatica, al livello del primo ordine, strutture classiche quale quella dei naturali (sulla cui non caratterizzabilità Skolem ritornerà più tardi in un fondamentale lavoro del 1933; si veda la nota 3 a pag. 562) quella dei reali ed infine quella stessa degli insiemi. Qui la difficoltà assumeva un carattere più preciso che coinvolgeva l'assolutezza delle nozioni insiemistiche, ad esempio quella di cardinalità. Per vederlo basta considerare il fatto seguente. All'interno di $\mathfrak{Z}\mathfrak{F}$ possiamo provare che esiste un insieme più che numerabile X, nel senso che X non è in corrispondenza biunivoca con i naturali; d'altra parte X conterrà solo elementi del dominio in cui ci muoviamo e quindi il dominio D non potrà essere numerabile. Ma il teorema di Skolem afferma che un modello numerabile deve esistere se $\mathfrak{Z}\mathfrak{F}$ è coerente: come risolvere questa situazione nota come *paradosso di Skolem*? Per Skolem, come abbiamo visto nel capitolo IV, l'alternativa è la seguente: o si passa a linguaggi di ordine superiore, accettando di fatto nella semantica l'univocità di concetti quale quello di parte la cui chiarificazione è compito della teoria degli insiemi, oppure accettiamo che proprietà insiemistiche come quella di numerabilità non sono assolute, nel senso che uno stesso insieme può essere numerabile in un modello e non numerabile in un altro. Non c'è nessuna contraddizione in questo, in quanto «numerabile» significa che esiste una biiezione con N che può trovarsi in un modello e non in un altro.

cardinalità $\mathfrak{m}$ di $\mathfrak{M}$ e maggiore o uguale alla cardinalità $\mathfrak{m}'$ dell'insieme delle costanti extralogiche di $\mathfrak{L}$. Come si può vedere, il teorema originale di Löwenheim-Skolem rientra nel secondo gruppo, mentre il primo risultato nella direzione i) si deve a Tarski e risale al 1934. Questo risultato si può ottenere dal teorema di completezza (o meglio da un suo corollario, il teorema di compattezza) dimostrato da Gödel per il caso numerabile e da Malcev per il caso generale. I limiti descrittivi dei linguaggi del primo ordine non riguardano quindi le sole sottostrutture ma anche le estensioni, che al pari delle prime possono essere indiscernibili dal punto di vista delle proprietà formulabili nel linguaggio. Come la ricerca successiva ha posto in luce, questi limiti non sono accidentali e non riguardano il solo linguaggio del primo ordine; teoremi analoghi si possono dimostrare anche per linguaggi più ricchi. Con alcune limitazioni, però: mentre i risultati di tipo ii) si estendono, poste le opportune modifiche, a linguaggi del tutto arbitrari vincolati alla sola condizione che la classe delle formule sia un insieme, i risultati di tipo i) si hanno solo per linguaggi che godano di una qualche forma del teorema di compattezza. Di questi temi avremo occasione di riparlare nel paragrafo 2 del prossimo capitolo, accennando agli sviluppi della teoria dei modelli.

Ma la storia dei metodi introdotti da Löwenheim non finisce qui. Qualche anno più tardi il francese Jacques Herbrand (1908-1931) nella citata dissertazione di laurea *Ricerche sulla teoria della dimostrazione* del 1930 dimostrava quella che può considerarsi la versione *finitista* del teorema di Skolem, a partire, per sua esplicita dichiarazione, dallo stesso ordine di idee del logico norvegese. Anche Herbrand non giungeva a compiere il passo finale che lo avrebbe portato al teorema di completezza e che avrebbe drasticamente semplificato la dimostrazione del suo teorema. Questa volta i motivi sono in qualche senso opposti a quelli di Skolem: è la professione di «fede» sintattico-finitista che non consente a Herbrand di porre il rapporto tra dimostrabilità e validità e per lui i concetti insiemistici sono privi di valore esplicativo sul piano epistemologico. Herbrand dimostrava, nel suo teorema, il seguente fatto: è possibile associare in modo effettivo, combinatorio, finitista, a ogni formula $\mathcal{A}$ del primo ordine una successione infinita di disgiunzioni della logica proposizionale, in modo tale che la formula $\mathcal{A}$ è dimostrabile nella logica dei predicati se e solo se una almeno delle disgiunzioni precedenti è dimostrabile nella logica degli enunciati. Se supponiamo cioè di avere una formula $\mathcal{A}$ del calcolo dei predicati le si può associare in modo effettivo una successione così fatta

$$
\begin{aligned}
H_1(\mathcal{A}) &= \mathcal{A}_1 \\
H_2(\mathcal{A}) &= \mathcal{A}_1 \vee \mathcal{A}_2 \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
H_i(\mathcal{A}) &= \mathcal{A}_1 \vee \mathcal{A}_2 \vee \ldots \vee \mathcal{A}_i \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}
$$

dove le $\mathcal{A}_i$ sono *formule proposizionali* (ottenute con metodo simile a quello skolemiano di approssimazione) e ogni H_i è detta appunto l'i-esima disgiunzione di Herbrand della formula $\mathcal{A}$. Questa associazione avviene in modo duale a quello visto per Skolem quando associavamo la formula $\mathcal{A}_S$ alla formula $\mathcal{A}$. Allora occorreva che $\mathcal{A}_S$ fosse *equisoddisfacibile* con $\mathcal{A}$; ora cerchiamo invece una formula $\mathcal{A}_H$ *equidimostrabile* con $\mathcal{A}$ in cui le costanti funzionali esemplificheranno i quantificatori universali e non quelli esistenziali. Consideriamo ancora la formula

$$\mathcal{A} \qquad \forall x \exists y \forall z \mathcal{B}(x, y, z);$$

per eliminare i due quantificatori universali introduciamo due costanti funzionali, una zeroaria, la costante individuale a, e una unaria h, e consideriamo la formula

$$\mathscr{A}_H \qquad \exists y \mathscr{B}(a, y, h(y))$$

(detta oggi *herbrandiana* di $\mathscr{A}$). Le varie disgiunzioni H_i non sono altro che le formule che si ottengono disgiungendo segmenti sempre più lunghi di una enumerazione degli esempi della matrice $\mathscr{B}$ di $\mathscr{A}_H$ ottenuti utilizzando i termini chiusi a, $h(a)$, ecc. e le loro combinazioni. Il collegamento con la tecnica di Skolem è immediato ove si pensi che una formula $\mathscr{A}$ è legge logica (la controparte semantica della dimostrabilità nella logica del primo ordine) se e solo se $\neg \mathscr{A}$ non è soddisfacibile, cioè se $(\neg \mathscr{A})_S$ non è soddisfacibile.

Il teorema allora afferma che $\mathscr{A}$ è dimostrabile nella logica dei predicati del primo ordine se e solo se esiste un n tale che $H_n(\mathscr{A})$ è un teorema della logica delle proposizioni e presenta forti analogie – nel senso che può svolgere lo stesso ruolo – col primo e secondo ε-teorema di Hilbert e Bernays. Siccome sappiamo, già dai lavori di Post, che tale logica è *decidibile*,[2] se la «riduzione» di Herbrand delle formule predicative a formule proposizionali fosse per così dire «totale» (basterebbe ovviamente che la successione di disgiunzioni per ogni formula fosse *finita*) avremmo automaticamente una procedura di decisione anche per la logica dei predicati. Vedremo più avanti (paragrafo 5) che così non è, ossia che la logica dei predicati *non* è decidibile in questo senso. Rimaneva comunque che la dimostrazione del teorema di Herbrand, diversamente da quella del teorema di Skolem, non richiede né l'assioma di scelta né nozioni infinitarie e proprio per questo è assai più complessa e informativa. Il suo grande significato sta nello stretto collegamento che stabilisce fra logica del primo ordine e logica proposizionale e nella possibilità che offre di stabilire confini numerici alle esemplificazioni esistenziali: in generale, nell'isolamento del significato costruttivo di teorie del primo ordine. Come tale fu utilizzato dallo stesso Herbrand, da Bernays e da altri ponendosi a fianco ai metodi di Ackermann basati sull'ε-calcolo. Diver-

[2] Ossia – come si ricorderà – che data una sua qualunque formula esiste una procedura meccanica che permette di decidere se essa è o no un teorema.

samente da quest'ultimo, però, soprattutto in anni recenti, le tecniche di Herbrand hanno rivelato un estremo interesse nello studio della dimostrazione automatica e in generale nell'informatica teorica.

Altra conseguenza del teorema di Herbrand che avremo occasione di richiamare in seguito, parlando di Gentzen, è che in base ad esso è possibile costruire per ogni formula $\mathscr{A}$ derivabile entro il calcolo dei predicati del primo ordine una «derivazione» che non faccia uso della regola del *modus ponens*. L'interesse della cosa sta nel fatto che, in questo modo, le dimostrazioni acquistano una forma particolarmente dominabile; con le parole di Herbrand «noi consideriamo questo fatto molto importante per delle difficoltà che si possono incontrare a causa della regola di implicazione [*modus ponens*] in certe dimostrazioni per ricorrenza [induzione]». Vedremo più avanti, considerando i lavori di Gentzen, come questo aspetto del teorema di Herbrand sia ricco di conseguenze per il programma hilbertiano e costituisca un passo importante verso l'analisi sintattica delle teorie.

Dal punto di vista del teorema di completezza, la situazione agli inizi degli anni trenta si può allora puntualizzare come segue secondo lo schema proposto da Jean van Heijenoort (1912-86). Si possono distinguere tre atteggiamenti fondamentali nella valutazione di una teoria: I. Il metodo *assiomatico*, in base al quale a partire da un insieme di assiomi si dimostrano formule («teoremi») sulla base di determinate regole di inferenza; II. Il metodo *semantico* che fonda la logica su nozioni contenutistiche di tipo insiemistico ed è interessato alla *validità* delle formule della teoria; infine III. Il metodo di Herbrand.

Consideriamo ora una formula qualunque $\mathscr{A}$, che per semplicità supporremo *chiusa*, di una teoria del primo ordine; a seconda che un autore condividesse o meno l'uno o l'altro dei tre atteggiamenti suddetti le questioni fondamentali circa $\mathscr{A}$ erano allora le seguenti. Un «assiomatico» si chiedeva: è questa formula dimostrabile oppure no? Un «insiemista»: è $\mathscr{A}$ valida oppure no? Herbrand infine: esiste un numero naturale n per cui l'n-esima disgiunzione di Herbrand della formula $\mathscr{A}$ risulti un teorema proposizionale oppure no? Il teorema di Herbrand afferma l'equivalenza delle «domande» assiomatiche con quelle da lui stesso poste; e Herbrand non si preoccupava di stabilire connessioni tra le sue esplicitazioni e le

questioni poste da un «semantico» perché «la nozione di validità si fonda su quella di sottoinsieme qualunque di un insieme arbitrario» e quindi non era ammissibile dal punto di vista finitista che egli professava; piuttosto Herbrand riteneva che tali nozioni dovessero essere sostituite dai concetti rigorosi da lui stesso introdotti. Il problema che rimaneva aperto era dunque quello dei rapporti fra una formula dimostrabile («herbrandiana» o no, vista l'equivalenza di cui sopra) e una formula valida: è proprio questo anello fondamentale della catena che il teorema di completezza semantica di Gödel sarebbe venuto a stabilire.

2. L'OPERA DI KURT GÖDEL FRA IL 1930 E IL 1940

È interessante notare come, libero da legami con una scuola particolare, sia stato proprio Gödel a trovare «immediato e naturale», una volta postosi dal punto di vista assiomatico, chiedersi se «il sistema inizialmente postulato di assiomi e di principi di inferenza è completo, ossia se esso effettivamente è sufficiente per la derivazione di ogni proposizione logico-matematica vera o se invece è concepibile che esistano proposizioni vere (addirittura eventualmente dimostrabili per mezzo di altri principi) che non possano essere derivate nel sistema considerato». La domanda era già stata posta da Hilbert e Ackermann nel 1928, ma tutto sommato non è certo che il senso nel quale essi attendevano una risposta fosse quello poi realizzato da Gödel. Quest'ultimo avrebbe risposto alla domanda dimostrando il teorema per cui *ogni formula valida del calcolo funzionale ristretto* [ancor oggi viene talora così chiamato il calcolo dei predicati del primo ordine] *è dimostrabile*, dopo averlo posto nella forma equivalente: *ogni formula del calcolo funzionale ristretto è o refutabile* [ossia è dimostrabile la sua negazione] *o soddisfacibile*. Egli osservava subito che l'equivalenza da lui dimostrata tra validità e dimostrabilità (nella logica dei predicati del primo ordine) «contiene per il problema della decisione una riduzione del non numerabile al numerabile, perché "valido" si riferisce alla totalità non numerabile di funzioni [interpretazioni possibili], mentre "dimostrabile" presuppone soltanto la totalità numerabile delle dimostrazioni formali». In altri termini, il teorema mostrava come una nozione come quella di validità veniva *simula-*

ta da una nozione più dominabile come quella di dimostrabilità.

Gödel non si limitava a questo e generalizzava quindi il teorema prima alla logica dei predicati con identità e poi a insiemi *arbitrari* al più numerabili di proposizioni del primo ordine: «Ogni insieme infinito numerabile di formule del calcolo funzionale ristretto è soddisfacibile (ossia tutte le formule del sistema sono simultaneamente soddisfacibili) o possiede un sottosistema finito il cui prodotto logico è refutabile». In altri termini, se agli assiomi logici aggiungiamo come nuovi assiomi extralogici un insieme infinito numerabile di proposizioni del primo ordine, la teoria che così ne risulta è ancora semanticamente completa.

La tecnica seguita da Gödel si basava essenzialmente su un approfondimento dei metodi di Skolem: si usano testimoni per esemplificare gli esistenziali e si costruisce il modello per successive estensioni utilizzando come dominio un opportuno insieme di termini. Per passare dalle soluzioni locali a quella globale che dà il modello Gödel fa esplicitamente riferimento ad «argomenti familiari» di tipo insiemistico che sostanzialmente possiamo assimilare al lemma di König sopra ricordato parlando di Skolem. Da allora numerosissime sono state le dimostrazioni di completezza per il calcolo classico di Hilbert-Bernays, come per altri calcoli, classici e non: indubbiamente sarà nota al lettore quella data da Leon Henkin nel 1949 e che fu poi estesa dallo stesso Henkin, nel 1950, alla teoria dei tipi (rispetto, si badi, a modelli generalizzati, sui quali si veda la nota 5 a pag. 567). È interessante osservare che anche questa dimostrazione prende le mosse da un lavoro di Gödel e precisamente da una breve comunicazione del 1932 dal titolo *Eine Eigenschaft der Realisierungen des Aussagenkalküls* (*Una proprietà delle realizzazioni del calcolo proposizionale*) in cui Gödel mostrava come ogni insieme non contraddittorio massimale dà origine a una valutazione che rende vere esattamente le formule dell'insieme. In questo contesto la costruzione dell'insieme massimale mostra forti analogie col teorema del filtro massimale per le algebre di Boole che a sua volta richiede l'assioma di scelta, anzi, come sarà provato da Henkin, è addirittura equivalente ad esso. La dimostrazione di Henkin estende il metodo alle teorie del primo ordine (e dei tipi) utilizzando l'aggiunzione di assiomi speciali che esemplificano gli esistenziali.

È allora chiaro che il concetto di completezza connette fra loro i concetti di verità e di dimostrabilità e non solo per la logica pura

ma per teorie arbitrarie. Assunto un calcolo come ad esempio quello di Hilbert-Ackermann, una teoria (l'aritmetica di Peano, $\mathfrak{Z}\mathfrak{F}$, ecc.) è data semplicemente specificando il suo insieme di assiomi extralogici o specifici, e teorema è ogni formula che si ottiene da questi assiomi assieme a quelli logici, utilizzando le regole logiche del calcolo. C'è uno stretto rapporto tra dimostrabilità entro una teoria e la proprietà di essere teorema logico, rapporto stabilito negli anni venti indipendentemente da Herbrand e da Tarski con il *teorema di deduzione*, il quale afferma che, limitandosi a formule chiuse, $\mathscr{A}$ è teorema di $\mathfrak{T} \cup \{\mathscr{B}\}$ sulla base del calcolo **K** ($\mathfrak{T} \cup \{\mathscr{B}\} \vdash_{\mathbf{K}} \mathscr{A}$) se e solo se $\mathfrak{T} \vdash_{\mathbf{K}} \mathscr{B} \rightarrow \mathscr{A}$. Iterando opportunamente l'applicazione del teorema e sfruttando il fatto che le dimostrazioni hanno lunghezza finita, si può concludere che $\mathfrak{T} \vdash_{\mathbf{K}} \mathscr{A}$ se e solo se esiste una successione finita $\mathscr{B}_1, \dots, \mathscr{B}_k$ di assiomi in $\mathfrak{T}$ tali che $\mathscr{B}_1 \wedge \dots \wedge \mathscr{B}_k \rightarrow \mathscr{A}$ è un teorema del calcolo **K**. Come si è detto sopra, oltre al teorema di completezza per il calcolo, Gödel provava anche che se un insieme T numerabile di formule è coerente, T possiede un modello. Questo ci permette di concludere che anche la nozione di *conseguenza*, limitata ad insiemi di assiomi numerabili, è simulabile in termini di dimostrabilità sulla base del calcolo. La dimostrazione che questo vale per ogni possibile insieme di formule anche *non numerabile* fu data successivamente nel 1941 da Anatolij Ivanovic Malcev (1909-67) seguito poi indipendentemente dallo stesso Henkin e da Abraham Robinson. Malcev poneva in luce un corollario fondamentale di questo teorema, la cui importanza Gödel non sottolineava. Si tratta del cosiddetto *teorema di compattezza*, in base al quale se $\mathscr{A}$ è conseguenza di $\mathfrak{T}$, $\mathscr{A}$ è conseguenza di una parte finita di $\mathfrak{T}$ o – in termini di soddisfacibilità – se ogni parte finita di $\mathfrak{T}$ ha un modello anche $\mathfrak{T}$ lo ha.

Come dice Mostowski, «il problema della completezza è un esempio interessante di una questione che sorge da ricerche filosofiche concernenti le relazioni tra calcoli e semantica e che ha trovato molte applicazioni puramente matematiche a dispetto della sua origine filosofica. Si parla spesso della rilevanza della logica matematica per l'algebra: è soprattutto il teorema di completezza che riesce a connettere queste due discipline in modo tale che esse possano profondamente e reciprocamente influenzarsi». Ciò risulta chiaro se si pensa al teorema di completezza espresso nella sua forma equivalente *ogni teoria noncontraddittoria del primo ordine ammet-*

te un modello, e si tiene conto che come modelli possiamo considerare le *strutture* di cui il matematico fa uso costante e naturale: insiemi non vuoti di elementi la cui natura non è specificata, e sui quali sono definite delle relazioni (strutture relazionali) o delle operazioni (strutture algebriche) o infine delle relazioni e delle operazioni (strutture). Queste applicazioni impiegano in generale il *teorema di compattezza* come sopra enunciato: un qualunque insieme infinito di formule chiuse del primo ordine ha un modello se ogni suo sottoinsieme finito ha un modello. Le prime applicazioni all'algebra di questo teorema risalgono a Malcev che lo impiegò nel 1941 in un lavoro dal titolo *Un metodo generale per ottenere teoremi locali nella teoria dei gruppi*, per risolvere e sistemare svariati problemi riguardanti rapporti tra proprietà dei sottogruppi e proprietà dei gruppi «ambiente», aprendo la strada a quello che sarebbe stato negli anni cinquanta lo studio dell'algebra dal punto di vista della teoria dei modelli.

Ritornando alla nostra storia, è naturale che la dimostrata completezza semantica delle teorie del primo ordine venisse immediatamente riguardata come un'ulteriore conferma della correttezza dell'impostazione hilbertiana nella ricerca sui fondamenti: l'apparato deduttivo formale era in grado di riprodurre fedelmente i rapporti semantici tra proposizioni e aveva dalla sua il vantaggio di farlo in termini finitisti, combinatori, senza ricorrere alle pesanti assunzioni insiemistiche che sono ineliminabilmente connesse con i concetti semantici. In altri termini, questo risultato sembrava offrire decisivi argomenti alla convinzione di poter *sostituire* ad ogni effetto il concetto di *dimostrabilità formale* a quello di *verità*. È però essenziale osservare, d'altra parte, che la posizione stessa del problema – se è possibile che le regole del calcolo siano adeguate a rendere la nozione insiemistica astratta di conseguenza – richiedeva di accettare come sensato il riferimento agli insiemi astratti. È questo che segna la rottura decisiva del lavoro di Gödel rispetto ai precedenti di Herbrand e Skolem. Gödel vede il problema come una questione che riguarda due ordini di nozioni che si presentano naturalmente nel pensiero astratto: nozioni insiemistiche non effettive, non dominabili, e nozioni combinatorie. Per Gödel – che esprime esplicitamente la sua posizione nella tesi ma non nel lavoro pubblicato – le nozioni infinitarie hanno senso, con buona pace di Hilbert, e la dimostrazione di completezza non mostra che esse

sono eliminabili, ma solo che esse sono *simulabili* se ci limitiamo al problema della conseguenza logica. Questo atteggiamento spiega come per Gödel sia naturale usare principi infinitari nella dimostrazione e paradossalmente è proprio per il suo distacco dalla ideologia finitista che Gödel ne fornisce una parziale conferma.

Ma altre ragioni c'erano per essere scettici sulla completa eliminabilità dei concetti infinitari. Già da tempo un limite dei linguaggi del primo ordine era stato posto in luce proprio da quel teorema di Löwenheim-Skolem che abbiamo visto nell'introduzione a questo paragrafo e che mostrava come ogni teoria del primo ordine con modelli infiniti fosse *non categorica*, ammettesse cioè modelli tra loro non isomorfi. Questo succedeva in particolare per la teoria degli insiemi e per la stessa aritmetica: per la prima lo stesso enunciato del teorema mostrava l'esistenza di modelli *numerabili* e quindi non certamente isomorfi al modello gerarchico che abbiamo descritto parlando di von Neumann o comunque ad un qualunque altro modello «intuitivo» di questa teoria, che intendiamo ovviamente come più che numerabile; per la seconda lo stesso Skolem costruirà un «modello non standard», cioè non isomorfo alla struttura N dei naturali, in quanto accanto ai numeri naturali contiene «altri» elementi e tuttavia soddisfa tutti gli ordinari assiomi dell'aritmetica (e quindi tutti i teoremi formalmente derivabili da essi[3]). In altri termini, i linguaggi del primo ordine pur se consentivano descrizioni complete delle teorie non riuscivano per così dire a «distinguere» i modelli delle teorie stesse, isolando solo quelli «desiderati». La descrizione di tali modelli non era cioè

[3] Già nel 1922, nell'articolo citato, Skolem aveva suggerito che con una data definizione della successione dei numeri naturali (grosso modo: alla maniera di Dedekind o di Frege) si ottenevano per essa modelli *diversi* nella teoria degli insiemi. E nel 1929 (nell'articolo citato) collegava più intrinsecamente la cosa alla teoria degli insiemi nel senso che dalle sue considerazioni risultava che la determinazione della successione dei numeri naturali dipendeva strettamente dagli assiomi scelti per la teoria degli insiemi. Nel 1933 infine in *Über die Unmöglichkeit einer Charakterisierung der Zahlenreihe mittels eines endliches Axiomensystems* (*Sulla impossibilità di caratterizzare la successione numerica per mezzo di un sistema finito di assiomi*) dimostrava il seguente teorema: Esiste un sistema N^* di cose, per le quali sono definite le operazioni $+$ e $\times$ e due relazioni $=$ e $<$, tale che N^* non è isomorfo al sistema N dei numeri naturali, ma cionondimeno tutte le proposizioni della teoria formale $\mathfrak{P}$ vere in N lo sono anche in N^*. Da questo teorema deriva esplicitamente il più debole corollario: nessun sistema di assiomi scritto nel linguaggio di $\mathfrak{P}$ può determinare univocamente la struttura della successione dei numeri naturali.

univoca cosicché se da un lato si disponeva di adeguatezza deduttiva (Gödel) dall'altro rimaneva un'ineliminabile inadeguatezza espressivo-descrittiva (Skolem).

A mostrare come i limiti della concezione formalistica fossero ancora più radicali, fu ancora una volta Gödel, con la celebre memoria *Über formal unentscheidbare Sätze der* Principia Mathematica *und verwandter Systeme*, I (*Sulle proposizioni formalmente indecidibili dei* Principia Mathematica *e di sistemi affini*, I) del 1931, che contiene quello che viene detto «teorema di Gödel» per antonomasia. Riferendosi in particolare a quelli che stima «i più comprensivi» sistemi assiomatico-formali, ossia quello dei *Principia* e la teoria degli insiemi (nell'assiomatizzazione di Zermelo-Fraenkel-Skolem o in quella di von Neumann) Gödel osservava: «si potrebbe quindi congetturare che questi assiomi e regole di inferenza siano sufficienti a decidere *ogni* questione matematica che possa essere formalmente espressa in questi sistemi. Verrà mostrato... che non è così, che al contrario esistono nei due sistemi menzionati problemi relativamente semplici della teoria dei numeri che non possono essere decisi sulla base degli assiomi». In altri termini Gödel dimostrava un teorema di *incompletezza sintattica* per l'aritmetica, che estendeva poi a una classe molto ampia di teorie; ma oltre a ciò egli dimostrava un risultato di *incompletezza semantica* per teorie basate su linguaggi di ordine superiore; infine derivava l'impossibilità di dimostrare la coerenza di un qualunque sistema (che soddisfi alcune naturali e importanti condizioni) nell'ambito del sistema stesso.

Per quanto riguarda il primo risultato, quello di incompletezza sintattica, Gödel riesce a costruire *in modo effettivo*, volendo porsi «al riparo da ogni critica intuizionista», una formula chiusa $\mathscr{G}$ del linguaggio della teoria dei numeri per la quale dimostra poi che tanto $\mathscr{G}$ quanto la sua negazione $\neg \mathscr{G}$ non sono derivabili nel sistema considerato, il quale quindi risulta appunto sintatticamente incompleto. (Per fissare le idee possiamo pensare di identificare il sistema entro cui Gödel si muove con la teoria $\mathfrak{P}$ (aritmetica di Peano) data nel paragrafo 3.4 del capitolo precedente).

Per questa dimostrazione Gödel si fonda su due idee fondamentali. La prima è quella dell'*aritmetizzazione*; essa consiste nell'associare a ogni simbolo dell'alfabeto, e quindi a ogni formula, successione di formule ecc. un numero naturale, detto appunto *numero di Gödel* dell'elemento linguistico considerato. Basterà associare ad

ogni simbolo dell'alfabeto, variabili escluse, un numero naturale diverso da 0; la scelta è del tutto arbitraria e riferendoci alla teoria $\mathfrak{P}$ data nel paragrafo 2.4 del capitolo precedente possiamo convenire ad esempio di fare la seguente assegnazione:

$$\begin{array}{ccccccccccccc} \bar{0} & s & + & \cdot & \vee & \neg & \wedge & \rightarrow & \exists & \forall & (&) & = \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 1 & 2 & 3 & 5 & 7 & 11 & 13 & 17 & 19 & 23 & 31 & 37 & 41 \end{array}$$

mentre alle variabili x_1, x_2, x_3,... assegniamo nell'ordine l'i-esimo numero primo a partire da 43. Dato il simbolo t indichiamo con $\ulcorner t \urcorner$ il corrispondente numero di Gödel. Per assegnare ora un numero di Gödel a ogni espressione (in particolare termini e formule) o successione di formule (come le dimostrazioni) ecc., si ricorre al teorema fondamentale dell'aritmetica per cui ogni numero naturale è scomponibile in uno e un sol modo (a parte l'ordine) in prodotto di potenze di primi. Porremo per definizione che data l'espressione $t_1...t_n$ il suo numero di Gödel sia

$$\ulcorner t_1...t_n \urcorner = p_1^{\ulcorner t_1 \urcorner} \times p_2^{\ulcorner t_2 \urcorner} \times ... \times p_n^{\ulcorner t_n \urcorner}$$

dove p_i è l'i-esimo numero primo.

Così ad esempio al termine $s(\bar{0})+ss(\bar{0})$ corrisponderà il numero

$$2^2 \times 3^{31} \times 5^1 \times 7^{37} \times 11^3 \times 13^2 \times 17^2 \times 19^{31} \times 23^1 \times 31^{37}.$$

Si noti che questa associazione è *effettiva* nel senso che: dato un qualunque elemento linguistico di $\mathfrak{P}$ sappiamo effettivamente calcolare il numero naturale che gli è associato; viceversa, dato un qualunque numero naturale sappiamo effettivamente decidere: 1) *se* esso è numero di Gödel di qualche elemento linguistico di $\mathfrak{P}$ e 2) in caso affermativo di *quale* elemento esso è numero. Ne risulta ora, in particolare, che le proprietà e relazioni sintattiche di $\mathfrak{P}$ (e quindi dei segni, delle formule, delle successioni di formule, ecc. di $\mathfrak{P}$) vengono automaticamente trasformate in proprietà e relazioni *numeriche*; non solo, ma si stabilisce in modo naturale una «reinterpretazione» delle proposizioni – che ora si riferiscono a numeri – in modo tale da «recuperare» il loro significato originario. In particolare, sarà possibile reinterpretare in questo nuovo «linguaggio aritmetico» quelle proposizioni del metalinguaggio che parlano di proposizioni del linguaggio, e in particolare di se stesse. Ora noi sappiamo che proposizioni che

parlano di proposizioni di un dato linguaggio L appartengono al metalinguaggio di L; e sappiamo anche che una confusione tra linguaggio e metalinguaggio comporta delle antinomie, tipica l'antinomia del mentitore, o quella di Richard ecc.

Orbene la seconda idea centrale di Gödel è proprio quella di riportare nella teoria, tramite l'aritmetizzazione, le proposizioni metateoriche sfruttando per così dire positivamente l'antinomia del mentitore per dimostrare l'incompletezza. Naturalmente, non è possibile, se si vuole conservare la coerenza della teoria, trovare in essa una proposizione (formula chiusa) che afferma di se stessa che non è vera, proprio perché altrimenti scatterebbe l'antinomia del mentitore; è però possibile, come Gödel mostra, costruire la formula $\mathscr{G}$ del linguaggio della teoria la quale, opportunamente reinterpretata via aritmetizzazione, afferma di se stessa (non che è non vera ma) che è *non dimostrabile* nella teoria. È proprio questa la formula che – come dicevamo – risulta in effetti non dimostrabile né refutabile e quindi è responsabile della incompletezza sintattica di $\mathfrak{P}$. Ci si può rendere conto intuitivamente della portata della sostituzione nel discorso relativo all'antinomia del mentitore, della nozione di *verità* con quella di *dimostrabilità*: nel primo caso, ossia per la verità, vale indubbiamente (almeno se si vedono le cose da un punto di vista classico) il terzo escluso, cioè una proposizione (una formula) deve risultare vera o falsa e sappiamo già che questo, applicato a una proposizione che esprime la sua propria falsità, porta ad antinomia, ossia rende contraddittoria la teoria cui quella proposizione appartiene; nel caso della dimostrabilità, invece, non necessariamente deve valere il terzo escluso, proprio perché non è detto che una proposizione debba necessariamente essere dimostrabile o refutabile in una teoria noncontraddittoria; può benissimo darsi il caso che non sia né l'una né l'altra cosa: la teoria cioè salva per così dire la propria coerenza sacrificando però la completezza sintattica. Da questo punto di vista il risultato di Gödel può essere presentato dicendo che una teoria (che naturalmente soddisfi le condizioni del teorema)[4] *deve*

[4] La teoria in questione dev'essere: 1) assiomatizzabile (ossia deve potersi decidere per ogni sua formula se essa è un assioma o no); 2) in grado di «esprimere» l'aritmetica (o almeno una certa parte di essa); 3) ω-coerente. La condizione 2), che può ritenersi banalmente verificata per $\mathfrak{P}$, può essere precisata in vari modi

essere incompleta sintatticamente pena la perdita di coerenza.

Si noti che il risultato di Gödel porta subito alla conclusione dell'esistenza di *modelli non standard*, non isomorfi a N, per l'aritmetica: e infatti poiché né $\mathscr{G}$ né $\neg\mathscr{G}$ sono dimostrabili in $\mathfrak{P}$, ognuna di queste formule può essere aggiunta agli assiomi di $\mathfrak{P}$ dando origine a due teorie, diciamo $\mathfrak{P}'$, $\mathfrak{P}''$ noncontraddittorie. Per il teorema di completezza ognuna di queste teorie ammette un modello (diciamo $\mathfrak{M}'$, $\mathfrak{M}''$ rispettivamente); ma $\mathfrak{M}'$ non può essere isomorfo a $\mathfrak{M}''$ perché in $\mathfrak{M}'$, poniamo, è vera $\mathscr{G}$, mentre in $\mathfrak{M}''$ è vera $\neg\mathscr{G}$. A differenza dei modelli non standard nominati parlando di Skolem, che risultavano per così dire indistinguibili dal punto di vista linguistico (cioè che sono, come oggi si dice, *elementarmente equivalenti*) il teorema di Gödel pone in evidenza una situazione ben più drastica, cioè di modelli non solo non isomorfi ma addirittura diversamente caratterizzati a livello linguistico.

Si vede anche facilmente come questo primo risultato di incompletezza sintattica si trasformi in un teorema di incompletezza semantica nel caso di teorie espresse in linguaggio di ordine superiore al primo. Consideriamo ancora il caso dell'aritmetica. Se si esprime questa teoria servendosi di un linguaggio del secondo ordine (nel quale, come si ricorderà, l'assioma di induzione mate-

(uno dei quali verrà accennato tra breve; si veda la nota 7 a pag. 569) e una teoria che la soddisfi viene oggi detta «sufficientemente potente»; la condizione 3) infine è una condizione più forte della semplice coerenza (ossia si possono avere teorie coerenti e tuttavia non ω-coerenti) e consiste sostanzialmente nel richiedere che se nella teoria $\mathfrak{T}$ si può dimostrare una data proprietà $P(x)$ *singolarmente per ogni cifra* (se si ha cioè $\vdash_{\mathfrak{T}} P(\bar{0}), \vdash_{\mathfrak{T}} P(\bar{1}), \vdash_{\mathfrak{T}} P(\bar{2})$ ... ecc.) non sia possibile nella teoria stessa dimostrare l'esistenza di un elemento che *non* goda di quella proprietà (non si possa cioè avere $\vdash_{\mathfrak{T}} \exists x \neg P(x)$). Quest'ultima condizione è stata ridotta alla semplice coerenza nel 1936 da J. Barkley Rosser in *Extensions of some theorems of Gödel and Church* (*Estensioni di alcuni teoremi di Gödel e Church*). È inoltre ovvio che devono essere «decidibili» le regole di inferenza assunte nella teoria, e questa è una condizione soddisfatta automaticamente in un sistema formale nel senso hilbertiano. Notiamo in proposito che Gödel stesso sottolinea il significato del suo teorema nei riguardi di più generali sistemi formali di cui però dichiara, nella memoria originale, di non possedere una precisazione rigorosa proprio per la mancanza di una adeguata definizione di regola «decidibile». In una rielaborazione della dimostrazione pubblicata nel 1934 afferma invece che grazie in particolare ai lavori di Turing sul concetto di decidibilità si dispone finalmente di tale definizione «inquestionabile»; si veda in proposito il paragrafo 6. Osserviamo infine che anche la prima condizione può essere indebolita come ha mostrato Mostowski nel 1952.

matica riceve la sua più naturale e completa espressione) come era già noto fin dal tempo di Dedekind, la teoria dei numeri diventa *categorica*, la teoria – come si dice – ammette uno e un solo modello a meno di isomorfismi. Orbene, costruiamo una formula (che chiameremo ancora $\mathscr{G}$) analoga a quella di Gödel nel senso che esprime la propria indimostrabilità ma rispetto alla situazione formulata al secondo ordine. Se ora supponiamo che $\mathscr{G}$ sia dimostrabile, essa sarebbe evidentemente falsa, sicché avremmo una proposizione dimostrabile e falsa, il che è impossibile in generale per la costruzione stessa delle nostre teorie (si ricordi quanto detto poco sopra); allora $\mathscr{G}$ non può essere dimostrabile, ma essendo proprio questo quanto $\mathscr{G}$ afferma, essa sarà vera. Abbiamo quindi una proposizione vera *nel* modello (e quindi in *tutti* i modelli, per la categoricità) ma non dimostrabile: la teoria è quindi *semanticamente incompleta*.[5]

A rigore, quello che abbiamo descritto è un risultato che vale solo per l'aritmetica del secondo ordine. Ma Gödel generalizza la cosa a *qualunque* teoria deduttiva espressa in un linguaggio di ordine superiore al primo e che abbia regole di inferenza abbastanza semplici da essere accettabili; fa allora vedere che una tale teoria è incompleta rispetto a una classe qualunque di modelli (purché almeno uno di questi modelli sia infinito).

Il secondo fondamentale risultato della memoria di Gödel del 1931 è il cosiddetto *secondo teorema di indecidibilità* (o secondo teore-

[5] La completezza semantica può essere «recuperata» per teorie (espresse in un linguaggio) di ordine superiore al primo ampliando la nozione di modello, come ha proposto Leon Henkin nel 1950. Naturalmente questo ampliamento può avvenire quando si disponga di una semantica rigorosa (e in particolare quindi di una precisa nozione di modello) per teorie del primo ordine, sicché la proposta di Henkin si fonda sui contributi in questo senso di Tarski, che noi vedremo nel prossimo paragrafo. È chiaro comunque che ad esempio al secondo ordine abbiamo nel linguaggio delle variabili (individuali) che vengono interpretate sugli *individui* del dominio del modello, e delle altre variabili (predicative) che vengono invece interpretate su *sottoinsiemi* del dominio stesso. Ora Henkin introduce per tali linguaggi il concetto di *modello generalizzato*, costituito da un modello nel senso di Tarski *con* una certa classe di sottoinsiemi del dominio del modello stesso; si impone allora alle variabili predicative la condizione di essere interpretate non su tutti i possibili sottoinsiemi del dominio, ma solo su quelli della classe sopraddetta. Si può far vedere che con questa nuova nozione di modello si ottiene, per teorie *arbitrarie* di ordine superiore al primo, un teorema di completezza analogo a quello stabilito da Gödel per le teorie del primo ordine.

ma di Gödel) che riguarda direttamente il programma formalista nella formulazione finitista hilbertiana. La dimostrazione di questo secondo risultato è solo schizzata nella memoria di Gödel e ne viene promessa una completa esplicitazione in un successivo lavoro che tuttavia non è stato mai pubblicato.

Abbiamo visto sopra che Gödel riesce, tramite l'aritmetizzazione, a «esprimere» all'interno di $\mathfrak{P}$ le relazioni e proprietà metateoriche di $\mathfrak{P}$ stessa; e qualunque cosa significhi quell'«esprimere» è ovvio che saranno delle *formule* della teoria formale $\mathfrak{P}$ che esprimeranno delle *proposizioni* della metateoria. In particolare, supponiamo che (almeno) una formula – che per comodità indicheremo con $Nctr_{\mathfrak{P}}$ – esprima in $\mathfrak{P}$ la proposizione metateorica «$\mathfrak{P}$ non è contraddittoria». Orbene Gödel fa vedere intanto che tale formula esiste e che in $\mathfrak{P}$ si può dimostrare l'implicazione $Nctr_{\mathfrak{P}} \rightarrow \mathscr{G}$, dove $\mathscr{G}$ è la formula indecidibile di Gödel. È allora chiaro che se in $\mathfrak{P}$ fosse dimostrabile la formula $Nctr_{\mathfrak{P}}$, per *modus ponens* (che è ovviamente una delle regole della teoria) si otterrebbe come teorema $\mathscr{G}$, contravvenendo però così al primo teorema di Gödel secondo il quale $\mathscr{G}$ è indecidibile. In altri termini: in una teoria sufficientemente potente e noncontraddittoria (in particolare, nell'aritmetica $\mathfrak{P}$) non è dimostrabile la noncontraddittorietà della teoria stessa. Se ricordiamo che Hilbert era giunto addirittura a identificare la metamatematica con l'aritmetica ricorsiva di Skolem, e che quindi *a fortiori* riteneva come sappiamo che ogni dimostrazione finitistico-combinatoria fosse formalizzabile nell'aritmetica, il secondo teorema di Gödel significa senz'altro la non realizzabilità di principio del programma di Hilbert di una dimostrazione finitista della coerenza dell'aritmetica.[6]

Il ruolo centrale giocato dalla formula $\mathscr{G}$, e la sua stessa costruzione e interpretazione pongono in particolare evidenza un mo-

[6] Conviene qui anticipare che l'enunciazione del teorema di Gödel sopra data non è del tutto adeguata in vista di alcuni risultati di Solomon Feferman in *Arithmetization of metamathematics in a general setting* (*Aritmetizzazione della metamatematica da un punto di vista generale*, 1960-1961). La situazione messa in luce da Feferman è *grosso modo* la seguente: siccome in generale esistono *più* formule di una teoria formale sufficientemente potente che possono esprimere formalmente *uno stesso concetto* o proprietà metamatematica, egli mostra come alcuni risultati – e in particolare proprio il secondo teorema di Gödel – possono dipendere per quanto riguarda la loro validità proprio dalla scelta effettuata, ossia dalla formula che si è assunta per esprimere un dato concetto, in questo caso la noncontraddittorietà.

mento fondamentale in tutto il lavoro di Gödel a cui ci siamo più volte riferiti nel nostro discorso: la possibilità di «rappresentare» nella teoria il suo aspetto metateorico, ossia le sue proprietà e relazioni metateoriche senza naturalmente cadere nelle note contraddizioni che sappiamo derivare dalla confusione fra linguaggio e metalinguaggio. Per comprendere chiaramente come Gödel riesca in questo suo intento, dopo aver dato una definizione di «rappresentabilità» in una teoria formale,[7] occorrerebbe richiamarsi alla teoria delle funzioni ricorsive cui noi accenneremo fra breve (si veda il paragrafo 6). Gödel assume alcune funzioni estremamente semplici e due schemi altrettanto intuitivi per generare nuove funzioni da funzioni date, col che individua tutta una classe di funzioni (oggi dette *ricorsive primitive*) che sembra di poter ragionevolmente assumere come corrispettivo rigoroso, anche se consapevolmente parziale, del concetto intuitivo di funzione effettivamente computabile (ossia tale che, comunque dati i suoi argomenti, sia sempre possibile, in un numero finito di passi, computare il valore corrispondente della funzione). Individua quindi le funzioni, proprietà e relazioni metateoriche interessanti la sua dimostrazione e fa vedere come queste risultino tutte ricorsive primitive; a questo punto dimostra in generale che sono «rappresentabili» nel senso precedentemente stabilito tutte le funzioni ricorsive primitive e

Egli definisce quindi in modo convenzionale una scelta «canonica» seguendo la quale risultati di questo tipo vengono conservati mentre costruisce addirittura un controesempio che fa vedere come il secondo teorema di Gödel non valga nel caso di una scelta non «canonica». La questione, di evidente interesse, è ancora oggetto di studio.

[7] Supponiamo per semplicità di avere l'aritmetica intuitiva da una parte e la teoria formale $\mathfrak{P}$ dell'aritmetica dall'altra. Ricordiamo che nel linguaggio di $\mathfrak{P}$ disponiamo in particolare di *nomi* per ogni numero naturale, ossia delle *cifre* $\bar{0}$, $\bar{1}$, $\bar{2}$, $\bar{3}$, ... Supponiamo ora di avere una relazione binaria $D(x, y)$ fra numeri naturali (si pensi ad esempio a $D(x, y)$ come a: x è divisibile per y). Gödel afferma che tale relazione dell'aritmetica intuitiva è *formalmente rappresentabile* in $\mathfrak{P}$ quando esiste una *formula* di $\mathfrak{P}$ con due variabili libere, sia $\mathscr{A}(x, y)$, tale che: ogniqualvolta la relazione intuitiva vale fra due numeri naturali a e b (si ha cioè $D(a, b)$) in $\mathfrak{P}$ si abbia $\vdash \mathscr{A}(\bar{a}, \bar{b})$ ossia sia dimostrabile la formula $\mathscr{A}$ nella quale al posto delle variabili libere siano state sostituite le *cifre* dei corrispondenti numeri; se viceversa D *non* vale fra due numeri c e d, in $\mathfrak{P}$ si deve avere $\vdash \neg \mathscr{A}(\bar{c}, \bar{d})$ ossia deve essere dimostrabile la negazione di $\mathscr{A}$ con l'analoga sostituzione delle variabili con le cifre. Questa definizione si estende in modo ovvio a relazioni n-arie qualunque come pure a funzioni n-argomentali qualunque.

quindi in particolare le relazioni metateoriche di cui sopra. Va anzi detto in proposito che quello di Gödel è il lavoro che, fra l'altro, ha dato un enorme impulso allo studio delle funzioni ricorsive, dopo i primi tentativi di Skolem cui abbiamo già fatto cenno (pagg. 528-530).

Notiamo ancora, per finire, che Gödel anche in questo lavoro come – in maniera più decisiva – in quello precedente dedicato alla completezza non si ritrae di fronte alle nozioni infinitarie della semantica, e ciò dichiara esplicitamente giustificando la sua posizione col dire che tali nozioni non dipendono da una particolare concezione della natura della matematica (e non possono quindi essere accettate o meno a seconda della «scuola» cui si aderisce) ma intervengono in modo naturale nella stessa pratica matematica quotidiana, sicché sarebbe erroneo non tenere conto di concetti così fondamentali e non «compromessi». Resta tuttavia da osservare che i riferimenti di Gödel a tali concetti sono ancora intuitivi, ossia la semantica cui egli si riferisce non ha ancora veste matematica. Questo è il compito che qualche anno più tardi assolverà Alfred Tarski, che già abbiamo avuto occasione di nominare, e la cui opera a questo riguardo prenderemo in esame subito dopo aver concluso questo excursus su Gödel.

Ci limiteremo solo ad accennare[8] a un terzo fondamentale risultato ottenuto da Gödel verso la fine degli anni trenta e che riguarda uno dei problemi centrali della teoria degli insiemi. Si tratta della dimostrazione della compatibilità o coerenza relativa dell'assioma di scelta (*AS*) e dell'ipotesi generalizzata del continuo (*IGC*) con gli altri assiomi della teoria su cui torneremo nel prossimo capitolo. Qui ci soffermeremo più in dettaglio sul senso logico delle tecniche usate da Gödel in questo contesto. Supposto che la teoria degli insiemi (intesa assiomatizzata in $\mathfrak{ZF}$ oppure mediante un sistema Σ di tipo $\mathfrak{NBG}$, in entrambi i casi senza l'assioma di

[8] Con ciò non esauriamo certo la presentazione dei suggerimenti e dei risultati basilari con i quali Gödel ha enormemente stimolato e indubbiamente determinato gran parte della ricerca logica successiva; e infatti lo incontreremo ancora nella trattazione si può dire di ogni tema centrale nella ricerca di questo e del successivo periodo. Si tenga conto infine che quando Gödel dimostra questo teorema ha ormai alle spalle tutta la sistemazione della semantica data da Tarski fra il 1933 e il 1936 con la conseguente chiarificazione del concetto di insieme definibile (si veda il paragrafo 2).

scelta; li indicheremo rispettivamente con Σ^* e $\mathfrak{ZF}^*$) sia non contraddittoria, Gödel dimostra che essa resta tale anche se si aggiungono come nuovi assiomi *AS* e/o *IGC*: se $\mathfrak{ZF}^*$ è coerente anche $\mathfrak{ZF}^*$ + *AS* + *IGC* è coerente. In altri termini, pur essendo assiomi molto «potenti» *AS* e *IGC* non generano altre contraddizioni che non siano eventualmente già presenti nella «più debole» teoria $\mathfrak{ZF}^*$ (o Σ^*). Al problema Gödel dedica tre lavori: il primo è del 1937, porta il titolo *The consistency of the axiom of choice and of generalized continuum hypothesis* (*La coerenza dell'assioma di scelta e dell'ipotesi generalizzata del continuo*) ed è una semplice comunicazione dei risultati, di circa due pagine; il secondo è del 1939: *Consistency-proof for the generalized continuum-hypothesis* (*Dimostrazione di coerenza dell'ipotesi generalizzata del continuo*) e in esso vengono schizzate, talora in modo informale, le dimostrazioni necessarie al risultato (anche se il metodo è da alcuni punti di vista diverso da quello poi impiegato nell'ultimo lavoro del 1940) e il riferimento è alla teoria $\mathfrak{ZF}^*$; il terzo lavoro infine è la monografia del 1940, *The consistency of the continuum hypothesis* (*La coerenza dell'ipotesi del continuo*), la quale contiene lo svolgimento completo della dimostrazione con riferimento al sistema Σ che abbiamo presentato nel paragrafo 3.2 del capitolo precedente.

Non è possibile qui rendere conto particolareggiatamente della dimostrazione; tenteremo tuttavia di illustrarne il significato con un minimo di bagaglio specifico che verrà a sua volta chiarito a livello puramente intuitivo (e quindi con inevitabili inesattezze). Cominciamo con il chiarirci il significato generale del teorema. Che una certa proposizione sia *compatibile* (coerente) con gli assiomi di una data teoria vuol semplicemente dire che non è possibile dimostrare nella teoria la negazione della proposizione stessa (o come anche si dice, non è possibile *refutare* la proposizione in questione nella teoria). Se infatti ciò accadesse, se cioè in una data teoria $\mathfrak{T}$ (del primo ordine) si riuscisse a dimostrare la negazione $\neg \mathscr{A}$ di una data proposizione $\mathscr{A}$, ciò comporterebbe che, essendo $\neg \mathscr{A}$ un *teorema* di $\mathfrak{T}$, essa dovrebbe risultare vera in *ogni* modello di $\mathfrak{T}$, o, equivalentemente, che in *ogni* modello di $\mathfrak{T}$ dovrebbe risultare falsa $\mathscr{A}$. Ma allora, per far vedere che una certa proposizione $\mathscr{A}$ non è refutabile nella teoria $\mathfrak{T}$ (ossia che non $\vdash_{\mathfrak{T}} \neg \mathscr{A}$) cioè che è compatibile con gli assiomi di $\mathfrak{T}$, sarà sufficiente trovare anche un *solo* modello di $\mathfrak{T}$ che verifichi $\mathscr{A}$ (naturalmente questo discorso va-

le per un numero qualunque di proposizioni o assiomi «aggiuntivi»). Orbene, quello che Gödel fa è proprio questo: trova un *modello* (che egli indica con Δ) per Σ^* nel quale sono verificati tanto *AS* quanto *IGC*. Per far questo egli introduce un metodo sintattico, detto dei modelli interni, che è divenuto standard nella teoria degli insiemi.

Tentiamo di renderci conto di come ciò avviene. In generale, una teoria $\mathfrak{T}$ «parla» di un certo universo di «cose» (in particolare $\mathfrak{ZF}$ o Σ «parlano» di un universo di insiemi (e classi) che possiamo identificare con la gerarchia cumulativa di von Neumann definita alla fine del paragrafo 3.2 del capitolo precedente).

Quantificando variabili individuali di $\mathfrak{T}$ – come sappiamo – noi facciamo riferimento alla *totalità* di queste «cose»; se viceversa «scegliamo» una «proprietà» in $\mathfrak{T}$ ossia una *formula* con una variabile libera che indichiamo ad esempio con $\mathscr{A}(x)$, noi *isoliamo* entro questo universo una classe particolare di oggetti, precisamente quelli che godono della proprietà su cui verrà interpretata $\mathscr{A}(x)$ (ad esempio se $\mathfrak{T}$ fosse la geometria euclidea e se si interpretasse $\mathscr{A}(x)$ come «x è un triangolo», isoleremmo con ciò fra tutte le configurazioni geometriche, la classe dei triangoli). In altri termini: una proprietà in $\mathfrak{T}$ isola un sottouniverso dell'universo del discorso in cui $\mathfrak{T}$ viene interpretata, di cui $\mathfrak{T}$ «parla». Ora, noi possiamo pensare di *relativizzare* i quantificatori in $\mathfrak{T}$ a questo sottouniverso determinato dalla proprietà $\mathscr{A}(x)$; basterà evidentemente che a ogni espressione della forma

$$\forall x Q(x) \ldots$$

noi sostituiamo un'espressione della forma

$$\forall x (\mathscr{A}(x) \rightarrow Q(x) \ldots$$

e che a ogni espressione della forma

$$\exists x Q(x) \ldots$$

sostituiamo un'espressione della forma

$$\exists x (\mathscr{A}(x) \wedge Q(x) \ldots$$

Così facendo limitiamo appunto l'ambito dei quantificatori al sottouniverso individuato dalla proprietà $\mathscr{A}(x)$. (Nel nostro esempio: le formule quantificate non si leggeranno più «per ogni x ...» o «esiste un x ...» ma «per ogni x, se x è un triangolo allora...» e «esiste un x tale che: x è un triangolo e...») Ad esempio la formula $\forall x \exists y \forall z \mathscr{B}$ (dove per comodità supponiamo $\mathscr{B}$ senza quantificatori) diventerebbe, come si dice, *relativizzata* ad $\mathscr{A}(x)$:

$$\forall x\ (\mathscr{A}(x) \rightarrow \exists y\ (\mathscr{A}(y) \wedge \forall z\ (\mathscr{A}(z) \rightarrow \mathscr{B}\ ))).$$

Supponiamo ora di avere due teorie $\mathfrak{T}_1$ e $\mathfrak{T}_2$ che abbiano lo stesso linguaggio. Potremo dire intuitivamente di aver «tradotto» $\mathfrak{T}_1$ in $\mathfrak{T}_2$ se «relativizziamo» tutti i quantificatori delle formule di $\mathfrak{T}_1$ rispetto a una certa proprietà $\mathscr{A}(x)$ di $\mathfrak{T}_2$ (e infatti le formule di $\mathfrak{T}_1$ «parleranno» ora del sottouniverso di $\mathfrak{T}_2$ individuato appunto da $\mathscr{A}(x)$). Può inoltre avvenire che noi riusciamo a eseguire questa relativizzazione in modo del tutto particolare, in modo cioè che riusciamo a trovare una speciale $\mathscr{A}(x)$ di $\mathfrak{T}_2$ tale che quelle particolari formule di $\mathfrak{T}_1$ che sono i suoi assiomi, una volta relativizzate rispetto ad $\mathscr{A}(x)$ diventino *teoremi* di $\mathfrak{T}_2$. In questo caso diciamo che $\mathfrak{T}_1$ è *relativamente interpretabile* in $\mathfrak{T}_2$ o che ammette un *modello sintattico* in $\mathfrak{T}_2$. (È facile convincersi che se $\mathfrak{T}_1$ può essere relativamente interpretata in $\mathfrak{T}_2$ allora se $\mathfrak{T}_1$ è contraddittoria tale sarà anche $\mathfrak{T}_2$). Questo varrà in particolare se $\mathfrak{T}_1$ e $\mathfrak{T}_2$ sono una stessa teoria $\mathfrak{T}$ e il modello così indotto dalla condizione relativizzatrice $\mathscr{A}(x)$ verrà detto, in modo naturale, *modello interno* della teoria $\mathfrak{T}$. Orbene Gödel riesce appunto a definire in una particolare proprietà $\mathscr{A}(x)$, «*x è costruibile*» che permette di interpretare relativamente $\Sigma^* + AS + IGC$ in Σ^*. Per quanto detto sopra, ciò significa che la proprietà in questione individua un sottouniverso nella struttura cumulativa dei tipi che inoltre risulta essere un modello, detto appunto *modello costruibile* (o dei *costruibili*) degli assiomi di Σ^* come pure di AS e di IGC: è questo il modello Δ di Gödel.

L'articolazione del procedimento descritto può essere *grosso modo* riassunta nei seguenti passi. Si indichi con $\mathscr{A}^L$ la relativizzazione di una formula $\mathscr{A}$ di Σ^* rispetto alla proprietà $L(x)$, «x è costruibile». Allora si tratta di dimostrare:

1) per ogni assioma $\mathscr{A}$ di Σ^*, $\vdash_{\Sigma^*} \mathscr{A}^L$; 2) $\vdash_{\Sigma^*} AS^L$; e 3) $\vdash_{\Sigma^*} \mathrm{IGC}^L$. Gödel non procede direttamente a dimostrare 1), 2), 3), ma introduce il cosiddetto *assioma di costruibilità* che afferma sostanzialmente che

ogni insieme è costruibile (ovvero: la classe totale V coincide con la classe dei costruibili L, in simboli, $V = L$). Egli dimostra quindi: la 1) sopra data per gli assiomi di Σ^*; 2′) $\vdash_{\Sigma^*} V = L \rightarrow AC$; 3′) $\vdash_{\Sigma^*} V = L \rightarrow IGC$; mostra infine che nel modello Δ vale la relativizzazione dell'assioma di costruibilità: i costruibili nell'universo di Σ^* sono anche i costruibili in Δ, ossia 4): $\vdash_{\Sigma^*} L(x) \rightarrow L(x)^L$. È chiaro che da 1), 2′), 3′), 4) discendono 1), 2), 3).

Non accenneremo neppure, in questa sede, al metodo di verifica delle varie proposizioni; vogliamo invece rendere intuitivamente il «contenuto» della proprietà di costruibilità (ossia, in altri termini, il modo con cui viene costituito Δ) ricollegandoci all'idea della ramificazione dei tipi. Si è visto che la gerarchia di von Neumann generava la struttura cumulativa dei tipi semplici; orbene Gödel ripercorre per così dire quella struttura a partire dall'insieme vuoto; ma ora a ogni passo successivo della costruzione non vengono presi *tutti* i sottosistemi dell'insieme ottenuto nel passo precedente bensì solo quelli che è possibile *definire predicativamente* in Σ^*, ossia quelli per i quali esiste in Σ^* una formula definitoria i cui eventuali quantificatori siano «relativizzati» a insiemi già ottenuti in passi precedenti. Naturalmente, ogniqualvolta si giunge a un ordinale limite si riunisce tutto quanto si è ottenuto sino a quel momento, ossia si costituisce un insieme che è null'altro che la riunione di tutti quelli ottenuti in passi precedenti (ossia per ordinali minori). Quindi si prosegue con lo stesso procedimento per tutti gli ordinali. Si noti: *tutti* gli ordinali, non solo quelli predicativamente definibili. È questo l'elemento che fa sì che i costruibili non siano un modello per la teoria predicativa dei tipi. Si ottiene così una struttura

$$L_0 \subseteq L_1, \subseteq \ldots \subseteq L_\omega \ldots L_{\omega+1} \subseteq \ldots \subseteq L_\alpha \ldots$$

che viene appunto detta struttura cumulativa ramificata dei tipi la cui riunione L è modello di $\mathfrak{Z}\mathfrak{F}$ e ciò mostra la non assolutezza del concetto di parte, in quanto per ogni α in generale ogni L_α è strettamente incluso in V_α.[9] Un insieme si dirà *costruibile* se appartiene

[9] Come detto, la struttura degli L_α è strettamente contenuta in quella degli V_α. Si pensi soltanto al fatto che, giunti al passo $\omega+1$, $V_{\omega+1}$ contiene *tutti* i sottoinsiemi di V_ω e quindi un numero più che numerabile di elementi; $L_{\omega+1}$ viceversa contiene, della potenza di L_ω, *soltanto* quegli elementi *definibili* che pertanto sono al più $\aleph_0$. Ovviamente è essenziale che nel nostro linguaggio si riesca a trovare una formula $\mathscr{A}(x)$ tale che $\mathscr{A}(a)$ se e solo se a è costruibile. È appunto questa la formula $L(x)$ relativizzatrice definita da Gödel.

a (*almeno*) uno degli L_α. I costruibili sono così gli insiemi predicativamente definibili una volta che si accetti l'iterazione transfinita fornita dagli ordinali cantoriani. La validità di *AS* in questo universo dipende appunto dalla possibilità di controllare ogni costruibile codificandolo per mezzo di un ordinale, il primo α tale che l'insieme appartiene a L_α, e la sua definizione in $L_{\alpha-1}$. Le definizioni sono ordinabili sicché è possibile, per ogni insieme di costruibili, disporre di un buon ordinamento così da verificare *AS*.

Torneremo più avanti sui costruibili; qui vogliamo sottolineare che il risultato così ottenuto da Gödel pur in tutta la sua enorme portata dava solo informazioni *parziali* sui rapporti di *AS* e *IGC* rispetto agli altri assiomi della teoria degli insiemi. Queste due proposizioni non si possono refutare in Σ^* (o $\mathfrak{ZF}$); ma si possono dimostrare? O non sono invece delle ipotesi indipendenti da quegli assiomi? Vedremo nel prossimo capitolo come sia stata data una risposta definita a questa domanda, che costituisce forse il maggior risultato degli anni sessanta e comunque senz'altro uno dei principali della ricerca logica da Cantor in poi.

3. LA SEMANTICA TARSKIANA

Come ricordato più volte, tra gli eventi più significativi degli anni trenta va posta la creazione della semantica per linguaggi formali che inaugura tutta una nuova stagione della ricerca metamatematica: lo studio, *non più su base strettamente finitista*, delle proprietà delle teorie, in particolare di quelle che riguardano la interrelazione fra espressioni linguistiche e oggetti astratti. Concetti quali quello di verità, legge logica, conseguenza, definibilità, ecc. cadono tutti in questo ambito e divengono strumenti concretamente utilizzabili dal punto di vista matematico.

Der Wahreitsbegriff in den formalisierten Sprachen (*Il concetto di verità nei linguaggi formalizzati*), il saggio in cui Tarski intraprendeva una precisazione rigorosa della semantica, apparve in polacco nel 1933 e in tedesco nel 1935 ed era inserito nel contesto di quella concezione della metamatematica che, come ricordato, è caratteristica della scuola polacca e di cui la semantica tarskiana sarebbe diventata una componente fondamentale. Come già per i risultati di Gödel non ci soffermeremo qui sui dettagli della costruzione di

Tarski che nelle sue linee generali sarà nota ad un lettore che abbia seguito un corso di logica; ci limiteremo quindi a tentare di mettere in evidenza le linee generali e i presupposti. Punto centrale della ricerca di Tarski è il riconoscimento che sostanzialmente tutte le nozioni semantiche fondamentali possono essere definite in termini del concetto di verità, considerato come *attributo di proposizioni*;[10] e che la ricerca, che quindi si impone, di una rigorosa definizione di questo concetto fondamentale, è a suo parere senza speranza condannata a fallire ove si voglia riferirla a proposizioni del linguaggio comune: donde l'altra specificazione che compare nel titolo del suo saggio.

Partiamo allora dal linguaggio comune e vediamo più da vicino alcune delle difficoltà che esso presenta e che si oppongono a una definizione rigorosa di verità, una volta che si cerchi di dare tale definizione in modo che sia «formalmente corretta» – ossia non porti a situazioni contraddittorie – e «materialmente adeguata» – ossia specifichi per *ogni* proposizione le condizioni sotto le quali essa può essere considerata vera.[11]

Forse la più antica definizione del concetto di verità si trova nella *Metafisica* di Aristotele: «Dire di ciò che è che non è, o di ciò che non è che è, è falso, mentre dire di ciò che è che è, o di ciò che non è che non è, è vero.» È a questa *spiegazione* più che definizione o al-

[10] E Tarski ovviamente è conscio, così facendo, di trattare il concetto di verità in senso «restrittivo», che cioè non tiene conto ad esempio di contesti psicologici o estetici nei quali viene impiegato il termine «vero». Ritiene però di cogliere con la sua impostazione quello che a suo parere potrebbe chiamarsi «il concetto logico di verità».

[11] Si noti a questo proposito che Tarski cerca appunto di specificare le *condizioni* sotto le quali una proposizione possa essere considerata vera e non dei *criteri* di verità che permettano in un qualunque modo di stabilire o di decidere se *di fatto* la proposizione in questione è vera. Dirà Tarski nel 1969: «Qualunque cosa possa ottenersi dalla costruzione di una definizione adeguata del concetto di verità per un linguaggio scientifico, una cosa è certa: la definizione non porta con sé un criterio pratico per decidere se una particolare proposizione di tale linguaggio sia vera o falsa (e invero questo non è affatto il suo scopo)... Alcuni filosofi ed epistemologi sono propensi a rifiutare ogni definizione che non fornisca un criterio per decidere, per ciascun oggetto particolare assegnato, se esso cada o no sotto il concetto definito. Nella metodologia delle scienze empiriche tale tendenza è rappresentata dall'operazionismo ed è condivisa anche da quei filosofi della matematica che appartengono alla scuola costruttivista; in ambedue i casi, tuttavia, solo una piccola minoranza di pensatori è di questa opinione.»

le sue numerose varianti tutte comunque riunite sotto la denominazione di teoria classica o *corrispondentistica* della verità, che Tarski si ispira.

In quest'ordine di idee, data una specifica proposizione, ad esempio «la calcite è un minerale», possiamo pensare di esprimere tali condizioni dicendo semplicemente: *l'enunciato «la calcite è un minerale» è vero se e solo se la calcite è un minerale.*

Si notino due cose: *a*) ha senso fare una distinzione fra proposizioni vere o false solo quando le proposizioni stesse convoglino una certa informazione su un certo ambito di realtà (*qualunque* esso sia) sicché il concetto di verità che noi stabiliamo è in effetti una *relazione* fra fatti linguistici (enunciati) e domini a cui essi si riferiscono; *b*) nella condizione relativa alla calcite può sembrare di aver espresso un puro e semplice truismo, qualcosa come «Socrate è Socrate» o un qualunque altro esempio del principio di identità. Che le cose non stiano così, ed è essenziale comprenderlo, è dimostrato dal fatto che nell'equivalenza sopra scritta a sinistra del «se e solo se» compare un enunciato *fra virgolette*, a destra un enunciato. Per chiarire il significato della cosa, immaginiamo di dare un nome, ad esempio *P*, all'enunciato «la calcite è un minerale». Allora l'equivalenza di cui sopra verrebbe scritta: *l'enunciato P è vero se e solo se la calcite è un minerale*, che non dà più luogo ad alcuna ambiguità, perché a sinistra dell'equivalenza appare *il nome* di un enunciato, mentre a destra appare *l'enunciato stesso.*[12] Orbene, se realizziamo chiaramente questa differenza e ci convinciamo dunque che una eventuale definizione come la precedente sarebbe «sensata», non vuota, potrebbe solo restarci il dubbio circa una sua sostanziale «banalità». Ma da questo dubbio ci liberiamo subito con le seguenti considerazioni.

Se potessimo scrivere un'equivalenza come la precedente (che diremo *equivalenza di tipo T*, da Tarski) per *ogni* enunciato del linguaggio naturale, avremmo evidentemente risolto il nostro problema; d'altra parte non tutti gli enunciati del linguaggio naturale hanno la forma semplice, atomica, come quello sopra esemplifica-

[12] Si noti che se riprovassimo ora a formulare la condizione, si avrebbe *l'enunciato P è vero se e solo se P*; si otterrebbe allora addirittura un non senso dal punto di vista grammaticale (infatti dopo il «se e solo se» deve figurare *un enunciato*, non il *nome* di un enunciato).

to. A partire da enunciati di questo tipo noi ne formiamo di assai più complessi servendoci ad esempio dei connettivi. Allora, e qui sta la non banalità della cosa: *a*) dovremmo poter scrivere un'equivalenza come la precedente per *ogni* enunciato naturale atomico; ma il numero di tali enunciati è *infinito*, il che di fatto rende impossibile già questo primo passo; *b*) concediamo pure di avere in qualche modo superato questa difficoltà: dovremmo allora dare delle «regole» per inferire, dalla verità di enunciati atomici, la verità di enunciati composti mediante i connettivi, il che presuppone che si sappia analizzare compiutamente *ogni* enunciato composto in termini dei rispettivi enunciati atomici componenti, o che si sappiano esplicitare *tutte* le regole in base alle quali noi operiamo tali composizioni. In altri termini, vista l'infinità dell'insieme degli enunciati e la possibilità di considerare enunciati sempre più complessi, noi penseremmo di stabilire delle condizioni di verità per gli enunciati atomici e quindi, sulla base della costruzione stessa degli enunciati composti, di inferire le condizioni di verità per questi ultimi a partire da quelle dei singoli componenti elementari. Orbene, se tentiamo di applicare in questo senso un procedimento rigoroso al linguaggio ordinario ci imbattiamo in difficoltà pressoché insormontabili per la stessa natura imprecisa delle regole grammaticali e sintattiche di formazione degli enunciati (imprecisa, si intende, almeno da un punto di vista logico-matematico).

Ma un'altra ben più grave difficoltà si oppone a questo procedimento. Abbiamo già più volte ricordato, fra le antinomie che derivano da una confusione fra linguaggio e metalinguaggio, quella del mentitore; anche nella costruzione di Tarski come già in quella di Gödel, questa antinomia gioca un ruolo decisivo, pur se in certo senso di segno opposto. Nel linguaggio comune ovviamente si hanno proposizioni che parlano di altre proposizioni (ad esempio: «la proposizione con cui comincia questo capoverso è breve») e in particolare che parlano di se stesse (ad esempio: «Questa proposizione è composta di sette parole», che è una proposizione vera). Fra molte altre tuttavia abbiamo in particolare anche una proposizione che afferma la falsità di se stessa: «Questa proposizione è falsa». Già sappiamo che tale proposizione conduce ad un'antinomia e tale antinomia può essere evitata solo mediante una rigorosa distinzione tra linguaggio e metalinguaggio: distinzione evidentemente impossibile nel

complesso del linguaggio naturale (*universalità* di tale linguaggio).

Sono questi i motivi principali che spinsero Tarski a costituire la sua definizione limitandosi a *linguaggi formalizzati* i quali offrono appunto almeno due vantaggi: 1) sono totalmente dominabili da un punto di vista sintattico; 2) permettono la rigorosa separazione di livelli che sembra fondamentale per superare le antinomie. È allora opportuno vedere come Tarski affronta e risolve il problema per i linguaggi formalizzati. Per comodità noi ci riferiremo a un linguaggio predicativo del primo ordine, tenendo presente che si tratterà, nella sostanza, di dare una definizione di verità che per così dire comprenda come casi particolari *tutte* le equivalenze di forma T riferite alle formule chiuse del nostro linguaggio; e che tale definizione dovrà far sì che a ognuna di tali formule spetti *almeno* uno e *al massimo* uno dei due valori di verità Vero o Falso.[13] Occorre ancora tuttavia fare una precisazione. Noi vogliamo definire la verità per *enunciati*, vale a dire per *formule chiuse* del nostro linguaggio formalizzato. Ora per le formule chiuse ci troviamo di fronte a una difficoltà analoga a quella constatata per il linguaggio naturale: le formule *chiuse* «composte» non sono formate a partire da formule *chiuse* «semplici» bensì, genericamente, da *formule*; allora noi non possiamo stabilire univocamente i passi di formazione per formule chiuse, mentre ovviamente ciò è possibile, per la loro stessa costruzione, per le formule in generale. Tarski allora pensa di sfruttare questa dominabilità strutturale dell'insieme delle formule definendo dapprima il concetto di *soddisfacibilità* per formule in generale dal quale poi ricava il concetto di verità per formule chiuse.

Per cominciare, posto che la verità è un concetto relazionale, rispetto a che cosa è vero un enunciato? La risposta di Tarski rompe con ogni perplessità di carattere finitario o con restrizioni di qualunque genere, immettendo esplicitamente la semantica nel contesto della teoria degli insiemi: la verità è riferita a interpretazioni e le interpretazioni sono costrutti insiemistici.

Dato allora un linguaggio $\mathscr{L}$ del primo ordine (che per comodità supponiamo abbia i soli connettivi di negazione e disgiunzione e il

[13] In altri termini, la definizione deve essere tale da permettere di *derivare* da essa tutte le equivalenze di tipo T e principi quali quello di non contraddizione (non possono essere entrambe vere due proposizioni una delle quali sia la negazione dell'altra) e del terzo escluso (due proposizioni siffatte non possono essere entrambe false).

solo quantificatore esistenziale), una *interpretazione* $\mathcal{I}$ di $\mathcal{L}$ sarà una coppia $<D, g>$ costituita da un dominio (non vuoto) D e da una funzione g che associa a ogni *costante* (individuale, predicativa o funzionale) di $\mathcal{L}$ una ben determinata entità del (o sul) dominio (rispettivamente: un individuo, una relazione fra individui, una operazione fra individui). Per quanto riguarda le variabili individuali esse saranno intese variare su D, cioè ad esse (a differenza che alle costanti individuali) g non assegna alcun individuo fissato del dominio.

Data ora una interpretazione $\mathcal{I}$, consideriamo tutte le successioni infinite di individui del dominio D[14] e indichiamo una generica di tali successioni con s; per ognuna di queste successioni si fissa in un qualche modo (ad esempio, definendo una funzione in corrispondenza ad ogni successione) l'associazione delle variabili agli individui della successione: ad esempio si può fare in modo di associare all'i-esima variabile x_i l'i-esimo individuo della successione considerata. Ciò posto, si tratta di definire il senso della frase: «la successione s soddisfa la formula $\mathcal{A}$ nell'interpretazione data». La definizione avviene ovviamente per induzione sfruttando il modo di costruzione delle formule. Inizieremo quindi con le formule atomiche, che avranno in generale, come sappiamo, la forma $P_i^k(x_1, x_2, \ldots, x_k)$. Diremo allora che la successione s soddisfa la formula $P_i^k(x_1, \ldots, x_k)$ se, essendo $\overline{P}$ la relazione su D che, nella interpretazione $\mathcal{I}$, g ha assegnato alla costante predicativa P_i^k, e $b_1, b_2, \ldots, b_k$ gli individui della successione s associati alle variabili $x_1, x_2, \ldots, x_k$, gli individui in questione stanno nella relazione $\overline{P}$ (cioè $< b_1, b_2, \ldots, b_k > \in \overline{P}$, ricordando che una relazione a k argomenti è un insieme di k-uple). Induttivamente, poi diremo che s soddisfa $\neg \mathcal{A}$ se e solo se s non soddisfa $\mathcal{A}$; e che s soddisfa $\mathcal{A} \vee \mathcal{B}$ se e solo se o s

[14] Questa, che può sembrare una grossa complicazione, si rivela in effetti una enorme semplificazione; si tenga sempre conto del fatto che in tutto questo discorso la difficoltà è proprio quella di adeguare rigorosamente concetti che abbiamo visto essere ampiamente usati si può dire «da sempre» in senso intuitivo. Va da sé che la complicazione è solo apparente nel senso che si riesce a dimostrare facilmente che la soddisfacibilità di una data formula $\mathcal{A}$ dipende, in una data interpretazione, non da tutti gli infiniti individui di una successione, ma solo da quelli (e sono certamente in numero finito) che «corrispondono» alle variabili libere contenute in $\mathcal{A}$. Si osservi infine che, dal momento che il dominio D è per ipotesi non vuoto, esiste sempre almeno una successione infinita di elementi di D. Nel caso banale che D contenga un solo individuo a, tale successione sarà proprio $s = \{a, a, a, \ldots\}$.

soddisfa $\mathscr{A}$ o s soddisfa $\mathscr{B}$. Per quanto riguarda i quantificatori (ci limitiamo, come detto, all'esistenziale) diremo che la successione s soddisfa la formula $\exists x_i \mathscr{A}$ se e solo se esiste una successione s' che differisce da s *al più* per l'i-esimo posto (e si noti che s è in particolare una tale successione) e s' soddisfa $\mathscr{A}$.[15]

A questo punto giungiamo alla desiderata definizione perché diremo che una formula $\mathscr{A}$ è *vera* in una data interpretazione $\mathscr{I} = \langle D, g \rangle$, e scriveremo $\mathscr{I} \models \mathscr{A}$, se *tutte* le successioni soddisfano $\mathscr{A}$ (*falsa* se *nessuna* la soddisfa). E diremo infine che una formula $\mathscr{A}$ è *logicamente vera*, $\models \mathscr{A}$, o *valida* se, comunque si scelga l'interpretazione, $\mathscr{A}$ risulta vera. Si osservi ancora che le nozioni qui definite sono proprio le intuitive e normali nozioni che ognuno di noi ha certamente «chiare»; la difficoltà sta nel darne una definizione rigorosa e matematicamente usufruibile e dominabile. Ad esempio, la nozione di *modello* di una teoria $\mathfrak{T}$ (di un insieme di proposizioni Γ) acquista ora un preciso significato: diremo che un'interpretazione $\mathscr{I}$ è *modello* di una teoria $\mathfrak{T}$ (dell'insieme Γ) se e solo se tutti

[15] Ribadiamo che quello delle successioni è un mero espediente tecnico di sistematizzazione (e infatti la definizione si può dire anche in modo diverso; noi ci siamo attenuti a questa per restare aderenti quanto più possibile al discorso di Tarski). La sostanza della questione sta nel fatto che con questo metodo si possono in qualche modo «dominare» gli individui del dominio D in modo tale da recuperare poi le intuitive accezioni semantiche. Supponiamo ad esempio di considerare un linguaggio con una costante relazionale P_1^2, e in esso la semplice formula atomica $\mathscr{A} = P_1^2(x_1, x_2)$. Sia $\mathscr{I} = \langle D, g \rangle$ una interpretazione con $D = \mathrm{N} = \{0, 1, 2, 3, \ldots\}$ e $g(P_1^2) =$ la relazione $\leqslant$ definita su N. Consideriamo ad esempio la successione $s = (7, 9, 21, 1, 5, \ldots)$. Per quanto stabilito, a x_1 corrisponderà 7, a x_2 corrisponderà 9. Sicché avremo che $\mathscr{A}$ diverrà sotto questa interpretazione $7 \leqslant 9$, ossia s soddisfa la formula in questione sotto l'interpretazione data. Mantenendo la stessa interpretazione e la stessa $\mathscr{A}$, la successione $s' = (12, 3, 5, 14, \ldots)$ *non* soddisfa la formula data perché questa ora risulta «tradotta» in $12 \leqslant 3$. In altri termini, $\mathscr{A}$ sarà soddisfatta in $\mathscr{I}$ da tutte e sole quelle successioni il cui primo elemento è minore o uguale al secondo elemento; in definitiva da tutte quelle coppie di numeri $\langle a, b \rangle$ con $a \leqslant b$. E la definizione è posta in modo tale che la soddisfacibilità di una formula da parte di una successione s in una data interpretazione $\mathscr{I}$ dipenda in effetti soltanto dalle variabili *libere* che essa contiene, ossia dai particolari valori che si danno a quelle variabili nell'interpretazione. Ne viene che in particolare la soddisfacibilità di una *formula chiusa* $\mathscr{A}$ *non* dipende dalla particolare successione scelta (in una data interpretazione) nel senso che o ogni successione soddisfa $\mathscr{A}$ o nessuna successione la soddisfa: per formule chiuse cioè *soddisfacibilità* e *verità* (in una data interpretazione) coincidono; il che, ancora, rende il fatto intuitivo che una proposizione (corrispettivo semantico intuitivo di una formula chiusa) è vera o falsa.

gli assiomi di $\mathfrak{T}$ (tutti gli elementi di Γ) sono *veri* in $\mathcal{I}$, cioè $\mathcal{I} \models \mathfrak{T}$. In particolare, si è ora in grado di precisare la vaga nozione di *conseguenza logica* di cui il concetto di *derivabilità logica* – che aveva già da tempo trovato (in particolare, abbiamo visto, con Hilbert) una sistemazione ineccepibile e soddisfacente – si poteva considerare come corrispettivo sintattico. Ora infatti diremo che una formula $\mathcal{B}$ è conseguenza logica di una formula $\mathcal{A}$ (o, in generale, di un insieme di formule Γ, e scriveremo $\Gamma \models \mathcal{B}$) se e solo se in ogni interpretazione, ogni successione che soddisfa $\mathcal{A}$ (o che soddisfa simultaneamente tutte le formule di Γ) soddisfa anche $\mathcal{B}$.

Ultimo concetto metamatematico che finalmente si può chiarire è quello di *insieme definibile*, centrale nei dibattiti sulla teoria degli insiemi. Data un'interpretazione $\mathcal{I}$ e un sottoinsieme X del suo dominio D, esso sarà definibile in $\mathcal{I}$ se esiste una formula $\mathcal{A}(x)$ tale che X coincide con l'insieme di tutti gli $a \in D$ per cui $\mathcal{I} \models \mathcal{A}(a)$. Vista la sua importanza, non deve stupire che le prime ricerche di Tarski che sfruttano le nozioni semantiche sopra introdotte riguardassero proprio il concetto di definibilità. Nel 1930, in collaborazione con Kuratowski, Tarski studiava gli insiemi definibili di numeri reali e, ponendo in luce i legami tra quantificatore esistenziale e l'operazione di proiezione π: $X \times X \rightarrow X$, stabiliva un preciso collegamento tra la teoria degli insiemi proiettivi e gli insiemi definibili entro linguaggi elementari che avrebbe avuto sviluppi decisivi nella teoria dei modelli e in quella della ricorsività.

Potrebbe ora sembrare che, dal momento che la definizione di verità (o di soddisfacibilità) per una teoria avviene nel metalinguaggio della teoria stessa, impiegando l'aritmetizzazione di Gödel si possano mettere le cose in modo tale che – come le nozioni sintattiche possono essere riportate senza contraddizioni nella teoria – lo stesso avvenga per le nozioni semantiche; orbene, questo non succede e, anzi, un famoso teorema di Tarski (dimostrato appunto nell'articolo del 1933) afferma che non è possibile definire in una teoria la nozione di verità per la teoria stessa senza cadere in contraddizione. Così ad esempio non è definibile al primo ordine l'insieme dei numeri di Gödel delle formule vere in N, insieme che, come vedremo, pur non essendo aritmetico risulta iperaritmetico. In generale il teorema di Tarski significa allora che per poter definire le nozioni semantiche non solo dobbiamo porci nel metalinguaggio, ma quest'ultimo deve essere essenzialmente più

ricco del linguaggio oggetto; in particolare deve far uso di un'ampia porzione di teoria degli insiemi. Se chiamiamo *sintatticamente chiuse* le teorie che possono esprimere la propria sintassi, e *semanticamente chiuse* le teorie che possono esprimere la propria semantica, i risultati di Gödel e Tarski ci dicono da una parte che è possibile avere teorie noncontraddittorie sintatticamente chiuse (le quali però risultano, come abbiamo visto, [sintatticamente] incomplete e in particolare non riescono a dominare un'importantissima proprietà sintattica, quella della noncontraddittorietà); dall'altra che è impossibile, pena il presentarsi dell'antinomia del mentitore, avere teorie semanticamente chiuse. Detto in altro modo: sappiamo che, via aritmetizzazione, possiamo assegnare a ogni formula (in particolare a ogni formula chiusa) di una data teoria un *numero*, precisamente il numero di Gödel di quella formula; possiamo allora considerare l'insieme G_1 dei numeri di Gödel dei *teoremi* della teoria e l'insieme T_1 dei numeri di Gödel delle proposizioni *vere* della stessa teoria. Il discorso precedente può allora riassumersi dicendo che i due insiemi G_1 e T_1 *non* coincidono e questo proprio perché mentre G_1 è «rappresentabile» nella teoria, tale non è T_1 (in quanto ogni formula $\mathscr{A}$ che lo rappresenta costituirebbe una sua definizione). Si noti che, come nel caso del teorema di Gödel, questa situazione si presenta per qualunque teoria «sufficientemente potente» e quindi in particolare anche per $\mathfrak{Z}\mathfrak{F}$. È questa una ulteriore conferma (ora stabilita con discorso rigoroso da entrambi i punti di vista, sintattico e semantico) dell'impossibilità in generale di rispecchiare adeguatamente, a livello sintattico, il momento semantico, dell'impossibilità cioè – presupposta in modo fondamentale ed essenziale dal programma hilbertiano – che la nozione (semantica) di *verità* venga *in toto* sostituita da quella (sintattica) di *dimostrabilità*. Questa impossibilità di traduzione aritmetica è d'altra parte la conferma del carattere irrimediabilmente infinitario e non effettivo delle nozioni semantiche che quindi *debbono* essere «trattate» in teorie particolarmente potenti, quale appunto la teoria degli insiemi.

In questa prospettiva la definizione di verità data da Tarski acquista, come si è detto, un significato centrale nel progetto di edificazione di quella «metamatematica» (o *metodologia delle scienze deduttive*, come Tarski la chiama) che come abbiamo visto nel capitolo II si poneva come alternativa a quella finitista hilbertiana.

Un aspetto delle ricerche di Tarski che dava speciale significato alla semantica in questo contesto, stabilendo uno stretto contatto tra logica, algebra e discipline connesse, stava nel fatto – che ben presto fu realizzato – che le strutture algebriche come i gruppi, gli anelli ecc. il cui studio sistematico costituiva il nucleo dell'algebra astratta, si potevano vedere come particolari interpretazioni di specifiche teorie formali. Si apriva così la strada a quello studio dell'algebra dal punto di vista della teoria dei modelli che sarebbe stato uno dei temi centrali della ricerca degli anni cinquanta e che fu di fatto, almeno da certi punti di vista, inaugurato dallo stesso Tarski a partire dal 1935, nei suoi studi sul calcolo dei sistemi.

Un ultimo cenno infine al problema della semantica dei linguaggi naturali. In certo senso si può tranquillamente affermare che il linguaggio naturale, il più potente mezzo linguistico di espressione di cui disponiamo, è semanticamente chiuso; da ciò deriva per il teorema di Tarski l'impossibilità di una definizione non contraddittoria del concetto di verità per il linguaggio stesso. Abbiamo detto «in un certo senso» perché al linguaggio naturale non si possono applicare direttamente e senza problemi le conclusioni di Tarski proprio perché non si tratta di un linguaggio «formalizzato». Tarski tuttavia affermava appunto l'impossibilità di stabilire una rigorosa semantica (teoria del significato) per il linguaggio naturale, sostenendo che quanto era possibile fare era al massimo questo: considerare linguaggi formalizzati che si discostassero «il meno possibile» dal linguaggio naturale e a questi linguaggi di approssimazione applicare la teoria sopra esposta. Si noti in particolare che un grosso ostacolo a tale applicazione nel caso del linguaggio naturale è costituito dalla componente intensionale propria a certi termini e contesti di questo linguaggio, non presente ovviamente nei linguaggi formalizzati fino ad ora da noi considerati, che sono costruiti e interpretati in modo rigorosamente estensionale. Come ricorderemo tuttavia più avanti (in VI, 5) in tempi recenti, la linguistica, avendo a disposizione semantiche più «articolate» di quella di Tarski, nata dalla necessità di fornire una controparte interpretativa soddisfacente per i linguaggi modali (o intuizionisti e più in generale intensionali) ha affrontato globalmente anche questo problema, seguendo quella che appunto viene detta la «linea di Tarski» o semantica denotazionale. Si noti ancora che anche nel caso di queste nuove semantiche la sistemazione di Tarski resta come sfon-

do generale, come schema e momento di motivazione, e si ritrova comunque come caso particolare di semantiche più generali.

4. LA «FORMALIZZAZIONE» DELLA LOGICA E DELLA MATEMATICA INTUIZIONISTE

In tutt'altro ordine di idee e in un certo senso in un mondo speculare, ci troviamo se ci volgiamo ad un altro dei grandi eventi della ricerca logica degli anni trenta, la sistemazione di Heyting dei concetti fondamentali della matematica intuizionista. Sappiamo già che Brouwer non ammetteva la formalizzazione e l'assiomatizzazione proprio per ribadire l'indipendenza della matematica intuizionista dalla logica e dal linguaggio formale sicché, come dice John Myhill, c'è da chiedersi se parlare di «formalizzazione» dell'intuizionismo non sia una *contradictio in adiecto*. Il fatto che Heyting intraprendesse questo compito non deve però far credere che egli avesse in proposito idee diverse da quelle di Brouwer, e non deve creare il fraintendimento di una diversa e mutata concezione intuizionista a riguardo. Abbiamo messo in luce in paragrafi precedenti qual è il senso e il significato che ebbe in generale questa operazione di «formalizzazione». Nel 1930, accingendosi a presentare il suo sistema, Heyting infatti ribadisce ancora che «la matematica intuizionista è un processo mentale e ogni linguaggio, incluso quello formalistico, è solo un ausilio per la comunicazione. È impossibile in linea di principio costruire un sistema di formule equivalente alla matematica intuizionista, poiché le possibilità di pensare non possono essere ridotte a un numero finito di regole costruite anticipatamente» anche se questa separazione fra linguaggio e matematica non va intesa in senso assoluto come risulta dalle seguenti parole del 1946: «L'intuizionista... non cerca il rigore nel linguaggio ma nello stesso pensiero matematico. Nello stesso tempo mi sembra contraddire la realtà supporre che la matematica intuizionista nella sua forma pura consista solo di costruzioni nel pensiero del singolo matematico, costruzioni che esistono indipendentemente l'una dall'altra e fra le quali il linguaggio pone una connessione, molto tenue... L'intuizionista usa quindi il linguaggio ordinario come il linguaggio simbolico in quanto ausilio alla memoria. Dobbiamo guardarci dall'immagine fittizia del ma-

tematico con la memoria perfetta che potrebbe lavorare senza l'aiuto del linguaggio. Nella ricerca matematica concreta il linguaggio è coinvolto in modo essenziale fin dall'inizio; la matematica come ci si presenta, convertita in espressioni linguistiche, non è preceduta da una fase completamente staccata dal linguaggio, ma è preceduta da una fase nella quale il ruolo del linguaggio è molto meno importante che nella comunicazione.» Ancora un'osservazione, per quanto riguarda in particolare la concezione di Heyting circa il metodo assiomatico, al quale egli ascrive due differenti funzioni: una funzione creativa, ad esempio nella teoria classica degli insiemi dove per così dire assicura l'esistenza di oggetti non garantita da alcuna costruzione; e una funzione descrittiva che si esplica nel suo impiego come sistemazione ed abbreviazione. Solo in quest'ultima accezione l'impiego del metodo assiomatico è legittimo anche nella matematica intuizionista.

Anche se oggi sono correnti altri sistemi più «trasparenti» per la logica intuizionista (e del resto ne vedremo uno noi stessi nelle prossime pagine) preferiamo dare il sistema logico originale di Heyting come presentato in *Die formalen Regeln der intuitionistischen Logik* (*Le regole formali della logica intuizionista*, 1930) e successivamente in *Die formalen Regeln der intuitionistischen Mathematik*, II, III, (*Le regole formali della matematica intuizionista*, 1930).

Per la logica proposizionale Heyting propone i seguenti schemi di assioma, con l'unica regola di separazione o *modus ponens*.

$$\frac{\vdash \mathscr{A}, \vdash \mathscr{A} \rightarrow \mathscr{B}}{\vdash \mathscr{B}}$$

H1: $\mathscr{A} \rightarrow (\mathscr{A} \wedge \mathscr{A})$

H2: $(\mathscr{A} \wedge \mathscr{B}) \rightarrow (\mathscr{B} \wedge \mathscr{A})$

H3: $(\mathscr{A} \rightarrow \mathscr{B}) \rightarrow ((\mathscr{A} \wedge \mathscr{C}) \rightarrow (\mathscr{B} \wedge \mathscr{C}))$

H4: $((\mathscr{A} \rightarrow \mathscr{B}) \wedge (\mathscr{B} \rightarrow \mathscr{C})) \rightarrow (\mathscr{A} \rightarrow \mathscr{C})$

H5: $\mathscr{B} \rightarrow (\mathscr{A} \rightarrow \mathscr{B})$

H6: $(\mathscr{A} \wedge (\mathscr{A} \rightarrow \mathscr{B})) \rightarrow \mathscr{B}$

H7: $\mathscr{A} \rightarrow (\mathscr{A} \vee \mathscr{B})$

H8: $(\mathscr{A} \vee \mathscr{B}) \rightarrow (\mathscr{B} \vee \mathscr{A})$

H9: $((\mathscr{A} \rightarrow \mathscr{C}) \wedge (\mathscr{B} \rightarrow \mathscr{C})) \rightarrow ((\mathscr{A} \vee \mathscr{B}) \rightarrow \mathscr{C})$

H10: $\neg \mathscr{A} \rightarrow (\mathscr{A} \rightarrow \mathscr{B})$

H11: $((\mathscr{A} \rightarrow \mathscr{B}) \rightarrow (\mathscr{A} \rightarrow \neg \mathscr{B})) \rightarrow \neg \mathscr{A}$

Per la logica predicativa, oltre agli schemi H1-H11, Heyting aggiunge i due schemi di assioma

H12: $\forall x \mathscr{A} \rightarrow \mathscr{A}(x/y)$

H13: $\mathscr{A}(x/y) \rightarrow \exists x \mathscr{A}(x)$

e le regole

$$\frac{\vdash \mathscr{A} \rightarrow \mathscr{B}\,(x/y)}{\vdash \mathscr{A} \rightarrow \forall x\, \mathscr{B}(x)} \quad \text{e} \quad \frac{\vdash \mathscr{A}\,(x/y) \rightarrow \mathscr{B}}{\vdash \exists x\, \mathscr{A}\,(x) \rightarrow \mathscr{B}}$$

sottoposte alle solite restrizioni sulle variabili.

Naturalmente occorre fare attenzione a non interpretare «classicamente» (cioè lungo le linee della semantica di Tarski) i connettivi e gli operatori logici che figurano nelle espressioni precedenti; di precisazioni del significato dei connettivi intuizionisti ne sono possibili diverse: ci soffermeremo qui sulle due principali, in termini di costruzioni e di problemi presentate rispettivamente da Heyting e Kolmogorov negli anni trenta. A proposito dell'interpretazione degli operatori logici, conviene comunque premettere un'osservazione generale di John Myhill secondo il quale «per comprendere l'intuizionismo è necessario aver ben chiaro fin dall'inizio che il processo di *spiegare* una nozione logica intuizionistica è diverso dallo stesso processo relativo a una nozione classica. Quest'ultima viene spiegata fornendo delle condizioni di *verità* per le proposizioni nelle quali essa interviene... la prima dando delle condizioni di *asseribilità*...».[16]

Inizialmente Heyting in *Sur la logique intuitionniste* (*Sulla logica intuizionista*) del 1930 e nel già ricordato *La fondazione intuizionista della matematica*, propone di interpretare una proposizione p come

[16] Assai chiarificatrice ci sembra anche la seguente immagine, sempre riportata da Myhill, che la accredita a Hao Wang: «Nel linguaggio di Wang... l'intuizionismo è una matematica del conoscere e la matematica classica è una matematica dell'essere. L'immagine è: in entrambe le concezioni il matematico ha di fronte

«l'intenzione di una costruzione matematica che deve soddisfare a condizioni determinate. La dimostrazione di una proposizione consiste allora nella realizzazione della costruzione che essa richiede: $p \rightarrow q$ rappresenta allora l'intenzione di una costruzione che, da ogni dimostrazione per p, conduce a una dimostrazione per q». In generale comunque, «una funzione logica è un metodo che trasforma ogni data asserzione in un'altra asserzione. La negazione è una tale funzione, il cui significato è stato descritto molto accuratamente da Becker, in accordo con Husserl. Secondo lui, la negazione è qualcosa che ha carattere positivo, precisamente l'intenzione di una contraddizione connessa con l'intenzione originale». E ancora «$p \vee q$ rappresenta l'intenzione che è soddisfatta se e solo se almeno una delle due intenzioni p e q è soddisfatta. La formula che esprime la legge del terzo escluso è $p \vee \neg p$. Per una data asserzione p, questa legge può essere asserita se e solo se o è stato dimostrato p o p è stato ridotto a una contraddizione. Quindi una dimostrazione per la legge del terzo escluso dovrebbe consistere in un metodo che permette, data una qualunque asserzione, o di dimostrarla o di dimostrare la sua negazione». Questa prima interpretazione comporta una distinzione fra l'asserzione «p» e l'asserzione «p è stato dimostrato», cui Heyting dedica alcune considerazioni nei due articoli sopra citati. Successivamente però egli abbandona questa interpretazione sostituendola con la considerazione delle formule logiche come esprimenti semplicemente *costruzioni*, col che ovviamente scompare la distinzione precedente, in quanto ora una proposizione p può essere *asserita* se e solo se si è in grado di realizzare la costruzione che essa esprime.

Un'altra interpretazione viene suggerita nel 1932 da Kolmogorov in *Zur Deutung der intuitionistischen Logik* (*Per l'interpretazione della logica intuizionista*) ove una proposizione viene intesa come esprimente un problema, sicché il calcolo ora diviene un *calcolo di problemi*. Considerando le due interpretazioni sopra esposte, si posso-

un blocco di sassi (proposizioni) che è occupato a dividere in vere e false, sicché in ogni momento si hanno tre blocchi chiamati "note-come-vere", "note-come-false", "(finora)-incognite". La differenza è che nella concezione classica, ma non in quella intuizionista, i sassi del terzo blocco sono tutti contrassegnati con V [vero] o F [falso] prima ancora della scelta, nozione questa che l'intuizionismo riguarda come teologica. Questa differenza concerne non solo il *contenuto* dell'intuizionismo, ma [anche] il *modo* col quale noi chiarifichiamo i suoi concetti».

no così riassumere l'interpretazione dei connettivi e le condizioni di asseribilità ($\vdash$) di una proposizione. Con $\mathcal{A}(x)$ si intenda una proprietà di cui gli enti di un certo tipo possono godere, rispettivamente un problema sensato per enti di un certo tipo. Avremo allora:

	Heyting	Kolmogorov
$\mathcal{A}$	una proposizione $\mathcal{A}$ esprime una costruzione (matematica)	una proposizione $\mathcal{A}$ esprime un problema matematico
$\vdash \mathcal{A}$	è stata eseguita la costruzione espressa da $\mathcal{A}$	è stato risolto il problema espresso da $\mathcal{A}$
$\vdash \neg \mathcal{A}$	supposta eseguita la costruzione espressa da $\mathcal{A}$ si può effettivamente costruire una contraddizione	supposto risolto il problema espresso da $\mathcal{A}$ si può effettivamente ottenere una contraddizione
$\vdash \mathcal{A} \wedge \mathcal{B}$	è stata eseguita tanto la costruzione espressa da $\mathcal{A}$ quanto quella espressa da $\mathcal{B}$	è stato risolto tanto il problema espresso da $\mathcal{A}$ quanto quello espresso da $\mathcal{B}$
$\vdash \mathcal{A} \vee \mathcal{B}$	è stata eseguita la costruzione espressa da $\mathcal{A}$ oppure quella espressa da $\mathcal{B}$	è stato risolto il problema espresso da $\mathcal{A}$ oppure quello espresso da $\mathcal{B}$
$\vdash \mathcal{A} \rightarrow \mathcal{B}$	si sa eseguire effettivamente la costruzione espressa da $\mathcal{B}$ ogniqualvolta si sa eseguire la costruzione espressa da $\mathcal{A}$	si sa effettivamente riportare la soluzione del problema espresso da $\mathcal{B}$ a quella del problema $\mathcal{A}$
$\vdash \exists x\, \mathcal{A}(x)$	si sa effettivamente indicare un individuo che gode della proprietà $\mathcal{A}(x)$	si sa effettivamente risolvere il problema $\mathcal{A}(x)$ per un dato individuo x
$\vdash \forall x\, \mathcal{A}(x)$	per ogni particolare elemento x si può far vedere che gode di $\mathcal{A}$	si può risolvere il problema $\mathcal{A}$ per ogni particolare elemento dato.

Si noti che contrariamente all'interpretazione classica dei connettivi (come funzioni di verità) e dei quantificatori (come funzioni da insiemi di valori di verità a valori di verità) con queste interpretazioni essi non sono interdefinibili ossia sono, come sappiamo, indipendenti (il che è stato dimostrato rigorosamente da McKinsey negli anni quaranta). Combinando le interpretazioni per la negazione e la disgiunzione si vede subito che il terzo

escluso non vale, in quanto esso verrebbe a significare o l'effettuabilità di ogni costruzione o equivalentemente la risolubilità effettiva di ogni problema matematico. Ora, dice Heyting, «affermare questo principio senza che esista un metodo generale di risoluzione per i problemi matematici non può essere giustificato se non riferendosi alla convinzione che la soluzione, se incognita, debba ciononostante essere definita in qualche modo; ma ciò significherebbe porre alla base delle considerazioni matematiche un principio filosofico, il che abbiamo riconosciuto... come inaccettabile». Analogamente le interpretazioni sopra date non consentono di accettare nella logica intuizionista l'implicazione $\neg\neg p \rightarrow p$ (mentre evidentemente vale il viceversa $p \rightarrow \neg\neg p$), sicché intuizionisticamente, a differenza di quanto avviene nella logica classica, non si ha l'equivalenza tra una proposizione e la sua doppia negazione. Interessante anche è notare che Brouwer aveva dimostrato la validità intuizionista della negazione della negazione del terzo escluso, ossia la formula $\neg\neg(p \vee \neg p)$ che è equivalente a $\neg(\neg p \rightarrow \neg\neg p)$, il che esprime il principio brouweriano dell'assurdità dell'assurdità del principio del terzo escluso. In altri termini questo principio non ha intuizionisticamente (come invece avviene nel caso classico) validità universale, ma *non* viene universalmente refutato. Volendolo esprimere in altro modo, non si può dare un problema dimostrabilmente insolubile: «quando ci si limiti ai problemi costruttivi», dice Heyting, «non se ne può dare uno che sia dimostrabilmente insolubile, ma l'ipotesi che tutti i problemi siano risolubili si dimostra ingiustificata».

Dopo la pubblicazione dei lavori di Heyting vennero dimostrati alcuni importanti teoremi sui rapporti fra logica e aritmetica classiche e intuizioniste, dove per Heyting l'aritmetica intuizionista $\mathfrak{AI}$ si ottiene da quella di Peano semplicemente adottando la logica intuizionista invece di quella classica. Per quanto riguarda la logica, è chiaro intanto che tutte le formule accettate dagli intuizionisti sono *valide* classicamente, quindi il calcolo logico intuizionista è certamente contenuto propriamente in quello classico. Un interessantissimo risultato veniva ottenuto nel 1933 da Gödel in *Zur intuitionistischen Arithmetik und Zahlentheorie* (*Sull'aritmetica e la teoria dei numeri intuizioniste*) in cui dimostrava che il calcolo proposizionale classico (**CPC**) è interpretabile nel calcolo propo-

sizionale intuizionista (**CPI**); risultato che, assieme al precedente, sembrerebbe poter far concludere per la coincidenza dei due sistemi. Le cose naturalmente non stanno così: nella (ovvia) dimostrazione **CPI** $\subseteq$ **CPC** si lasciano inalterati i connettivi, sicché la traduzione delle formule intuizioniste in quelle classiche è *identica*: se $\mathscr{A}$ vale intuizionisticamente, allora $\mathscr{A}$ vale anche classicamente. È chiaro tuttavia che questa traduzione non è per così dire «fedele» (dal punto di vista intuizionista) proprio per le diverse interpretazioni che nei due casi ricevono i connettivi. L'implicazione inversa, **CPC** $\subseteq$ **CPI** dimostrata da Gödel è invece *fedele* in questo senso, cioè rappresenta la logica classica su quella intuizionista traducendo opportunamente le formule valide della prima in formule valide della seconda mediante una «traduzione» delle stesse che non «tocca» i significati classici dei connettivi, pure interpretandoli correttamente dal punto di vista intuizionista. Sicché ora, data una formula $\mathscr{A}$ in **CPC**, Gödel le associa una formula $\mathscr{A}^*$ (in generale *diversa* da $\mathscr{A}$) valida in **CPI**. Il passaggio avviene traducendo le nozioni classiche $\neg p$, $p \rightarrow q$, $p \vee q$, $p \wedge q$ nelle corrispondenti nozioni intuizioniste $\neg p$, $\neg(p \wedge \neg q)$, $\neg(\neg p \wedge \neg q)$, $p \wedge q$.[17] Si può estendere il risultato al calcolo dei predicati traducendo il quantificatore esistenziale $\exists x$ in $\neg \forall x \neg$ (lasciando invariato l'universale). Ora Gödel dimostra che la stessa relazione vale fra l'aritmetica di Peano basata sulla logica classica (Gödel si riferisce a una variante di un sistema dato da Herbrand nel 1931) e quella basata sulla logica intuizionista: anche in questo caso si possono interpretare le nozioni classiche in termini di nozioni intuizioniste in modo che «tutti gli assiomi classici diventino proposizioni dimostrabili anche per l'intuizionismo»; e anche ora si tratta di associare opportunamente ad ogni formula classica $\mathscr{A}$ la sua traduzione $\mathscr{A}^*$, in modo

[17] Ad esempio la tautologia classica $\mathscr{A} = p \rightarrow (q \rightarrow p)$ diventerebbe $\mathscr{A}^* = \neg(p \wedge \neg \neg(q \wedge \neg p))$ e $\mathscr{A}^*$ risulta ora valida intuizionisticamente; se $\mathscr{A} = p \vee \neg p$ ossia il terzo escluso, si avrà $\mathscr{A}^* = \neg(\neg p \wedge \neg \neg p)$ e così via. Si comprende meglio il senso di tutto il discorso se si pensa che la traduzione identica (da **CPI** a **CPC**) *non* può essere intuizionisticamente adeguata dal momento che l'interpretazione intuizionista dei connettivi richiede ad esempio una distinzione fra le due formule p e $\neg \neg p$ che invece risultano, come sappiamo, classicamente equivalenti. È chiaro allora come avendo viceversa a disposizione questa analisi più «sottile» si possa operare in modo fedele il passaggio inverso.

tale che se $\mathscr{A}$ è dimostrabile nell'aritmetica classica, $\mathscr{A}^*$ sia dimostrabile nell'aritmetica di Heyting.[18]

Sostanzialmente una traduzione simile (limitatamente alla logica) era già stata data da Kolmogorov nei lavori citati, e interpretazioni di questo tipo (note come interpretazioni *negative*) furono presentate più tardi da Gentzen e nel 1951 da Kuroda. Connessi a questi lavori sono alcuni precedenti risultati (che risalgono al '28 e '29) di V. Glivenko che provò che per la logica proposizionale si ha

$$\vdash_{\mathrm{K}} \mathscr{A} \Leftrightarrow \vdash_{\mathrm{I}} \neg\neg\mathscr{A}$$

dove $\vdash_{\mathrm{I}}$ va letto «dimostrabile intuizionisticamente». Questo risultato di Glivenko è importante perché fornì una prima misura di come la logica classica fosse un rafforzamento di quella intuizionista: è infatti immediato concludere da quanto sopra che la prima si ottiene dalla seconda aggiungendo lo schema del terzo escluso per cui $\neg\neg\mathscr{A} \rightarrow \mathscr{A}$. Altra conseguenza è che per ogni formula «negativa» $\neg\mathscr{A}$ dimostrabilità classica e intuizionista (proposizionali) coincidono in quanto già Brouwer aveva dimostrato che $\neg\neg\neg\mathscr{A} \leftrightarrow \neg\mathscr{A}$.

Quale fosse il significato di questi risultati è Gödel ancora a renderlo esplicito nell'articolo citato.

«La dimostrazione qui effettuata», afferma Gödel, «mostra che *l'aritmetica e la teoria dei numeri intuizioniste sono solo apparentemente più ristrette delle versioni classiche, e in effetti le contengono (impiegando un'interpretazione alquanto deviante)*. La ragione di ciò sta nel fatto che la proibizione intuizionista di negare proposizioni universali per formare proposizioni puramente esistenziali viene resa inoperante ammettendo l'applicazione del predicato di assurdità alle proposizioni universali, il che formalmente conduce esattamente alle stesse proposizioni come sono asserite nella matematica classica. L'intuizionismo sembrerebbe risultare una genuina restrizione so-

[18] Questo risultato dà evidentemente una dimostrazione di coerenza intuizionistica per l'aritmetica «classica» di Peano. «È notevole», osserva Mostowski, «che questa dimostrazione risulti essere così semplice mentre non esistono dimostrazioni di coerenza strettamente finitiste. Quindi l'aritmetica di Peano basata sulla logica intuizionista contiene molti elementi non finitistici».

lo per l'Analisi o la teoria degli insiemi e queste restrizioni sono il risultato non della negazione del tertium non datur, ma piuttosto della proibizione di concetti impredicativi. Le considerazioni precedenti danno ovviamente una dimostrazione di coerenza per l'aritmetica e la teoria dei numeri classica. Tuttavia questa dimostrazione è certamente non "finitaria" nel senso dato da Herbrand seguendo Hilbert». Ben lungi dall'essere un indebolimento della logica classica, quella intuizionista si rivelava un suo raffinamento.

Al di là del significato fondazionale, risultati di collegamento come quello di sopra davano indicazioni su come affrontare un problema tecnico di grande importanza: quello di elaborare criteri maneggevoli (nel caso ottimale, metodi di decisione) per determinare la dimostrabilità o meno di date formule all'interno di teorie intuizioniste.

In questa prospettiva S. Jaśkowski in *Recherches sur le système de la logique intuitionniste* (*Ricerche sul sistema della logica intuizionista*, 1936) dimostrava un teorema che riporta la dimostrabilità di una formula intuizionista proposizionale alla sua validità in opportune matrici; come abbiamo ricordato, nel 1938 Tarski in *Der Aussagenkalkül und die Topologie* (*Il calcolo proposizionale e la topologia*) impiegò questo teorema per dare un'interpretazione topologica del calcolo proposizionale intuizionista ottenendo così un teorema di completezza rispetto a interpretazioni in algebre di aperti.

Si dovranno attendere gli anni sessanta per avere analisi semantiche più significative della logica intuizionista dal punto di vista intuitivo, analisi che col passare del tempo perderanno il loro aspetto strumentale per rivelare invece come, in modo naturale, all'interno della matematica classica si presentino situazioni analizzabili in termini intuizionisti e viceversa distinzioni intuizioniste siano simulabili classicamente.

Limitarsi alla logica, però, non può dare un'idea adeguata di come, al di là delle dichiarazioni di principio, si possa sviluppare la *matematica* intuizionista e di quale sia in questo sviluppo il ruolo della logica. A partire dalla fine degli anni venti le idee di Brouwer si cristallizzarono gettando le basi di quella matematica intuizionista che Brouwer contrappone alla classica. È su questo sfondo che va valutato il rigetto della logica classica da parte di Brouwer, in quanto è proprio la consapevolezza che c'è un nesso tra tipi di oggetti che si considerano e la logica che si utilizza a separare la

posizione intuizionista da quelle predicativiste e genericamente costruttiviste di Poincaré, Borel, Lebesgue, ecc. Abbiamo già ricordato come per Brouwer le costruzioni mentali vengano prima del linguaggio e come la logica non sia altro che la teoria (empirica) delle regolarità del linguaggio, in particolare di quelle regolarità di cui si occupa la sillogistica e che hanno a che fare con l'intero e le parti. Sulle prime le conseguenze che Brouwer trasse da queste premesse furono molto diverse da quelle che avrebbero poi segnato il sorgere della matematica intuizionista. «La matematica del tutto e della parte», osserva infatti Brouwer in una lettera al suo maestro Korteweg nel 1907, «come teoria non ci insegna nulla di nuovo per quanto riguarda la pratica: una volta che il sistema è stato applicato a una parte del mondo della visione interna, un intelletto anche mediocre può trarvi tutte le conseguenze e non c'è bisogno di ragionamento logico intermedio».

Sostanzialmente quindi, almeno all'inizio, Brouwer più che rigettare la logica classica sembra negare *tout-court* un ruolo significativo al ragionamento logico-ipotetico. È solo lentamente, a partire dal 1908, che si viene a configurare quella sostanziale intertraducibilità del terzo escluso col principio di risolubilità di ogni problema (nel senso di Hilbert) che rimarrà una costante della critica intuizionista. Questo avviene in modo pieno quando finalmente Brouwer spezza i legami con i costruttivisti come Borel e Lebesgue e passa a considerare oggetti di tipo nuovo per la matematica: le *successioni di (libera) scelta*. A questo punto non solo è possibile una ricostruzione aritmetica del continuo geometrico e quindi un pieno sviluppo della matematica intuizionista, ma il ragionamento ipotetico acquista un ruolo centrale nel pensiero matematico, con la conseguente necessità del rigetto non più della logica in sé, ma della logica *classica*, a cominciare dal terzo escluso.

È questo un punto su cui Brouwer non avrà più ripensamenti. La ricostruzione aritmetica del continuo permetteva – come vedremo – di fare a meno di un'intuizione fisica o geometrica dello spazio come ancora era per Poincaré e Borel e per lo stesso Brouwer della Tesi. Questo avrebbe però comportato l'abbandono della logica classica e l'accettazione del ragionamento ipotetico e della irriducibilità delle proprietà negative. In interessantissimi scambi epistolari tra Heyting e H. Freudenthal e tra Heyting e Kolmogorov, attorno ai primi anni trenta, ci è dato verificare come ci fosse

tutta un'ala estrema – rappresentata da matematici come lo stesso Freudenthal e G.F.C. Griss, van Dantzig e Vredenduin – per cui non era accettabile la necessità di costruzioni ipotetiche e non attuali (come quelle richieste per giustificare le implicazioni) e di proprietà negative (che lette in termini costruttivi coincidono con costruzioni di proprietà irrealizzabili). Lungo questa linea, a partire dal 1946 Griss avrebbe sviluppato una sua matematica intuizionista *senza negazione*, mentre nel 1948 Brouwer mostrerà proprio contro di lui come esistano proprietà *essenzialmente negative*, che non si possono definire senza la negazione.

Se per mostrare la non affidabilità della logica classica già dal 1908 Brouwer si era servito dei cosiddetti *controesempi deboli* in cui la costruzione di particolari reali o successioni di reali mostrava come la validità di teoremi classici costruttivamente intesa comportasse la risoluzione di problemi non risolti, la necessità di una nuova logica, almeno a partire dal 1918, è per Brouwer legata – come sopra abbiamo detto – all'introduzione delle successioni di scelta.

È un fatto ovvio che qualsiasi concezione della matematica che pretenda di avere un minimo di credibilità non può evitare di fornire un'analisi del continuo, che è alla base della matematica moderna e la difficoltà centrale che tutti i costruttivisti fino ad allora avevano trovato era stata appunto quella di conciliare il carattere discreto e numerabile delle possibili definizioni costruttivamente significative dei numeri reali, con la continuità e la densità che l'intuizione fisica e geometrica mostrano essere caratteristiche irrinunciabili dello spazio. Anche senza accettare in positivo la dimostrazione cantoriana della più che numerabilità del continuo, era inevitabile la conclusione che esso almeno *non* era numerabile, mentre di necessità numerabile doveva essere la totalità delle definizioni di numeri, di funzioni, ecc. in termini finiti che i costruttivisti richiedevano.

A questo si aggiungano da una parte le difficoltà connesse con il concetto di definibilità messo in luce dall'antinomia di Richard (con la conseguente necessità puntualizzata da Poincaré di limitarsi a definizioni predicative) dall'altra il fatto che in base alle diverse teorie della misura sviluppate (da Jordan, Borel, Lebesgue, ecc.) l'intervallo chiuso [0, 1], una volta ridotto ai suoi punti definibili, sarebbe risultato di misura zero, vale a dire irrilevante dal

punto di vista della misura. Questo spiega la sostanziale ambiguità dei costruttivisti francesi che in pratica continuarono a fare matematica come gli altri, tutt'al più limitandosi a scegliere i temi di ricerca in conformità ai propri presupposti ideologici, ma rinunciando a sviluppare un vero *corpus* organico di matematica costruttiva come a suo tempo – ma in ambito più ristretto – aveva tentato di fare Kronecker, cercando di dar forma al suo *Jugendtraum* di una algebro-geometria su base finitista.

Fra tutti, è forse Borel che nelle sue oscillazioni ci permette di capir meglio il senso della strada presa da Brouwer, anticipandone l'introduzione delle successioni di scelta. Borel è consapevole dell'abisso che esiste fra il continuo aritmetico che il costruttivista è in grado di ottenere e il continuo geometrico intuitivo: mentre quest'ultimo ci è dato *in toto*, come una struttura globale, in termini geometrici per cui – in un certo senso – prima viene la totalità e poi i punti che la compongono, il continuo aritmetico viene costruito come insieme di punti, e i punti sono i numeri reali che esistono *prima e indipendentemente* dalla totalità di cui fanno parte. Per un costruttivista, però, una definizione dei reali come quella di Cantor che fa riferimento a successioni di Cauchy arbitrarie, non può avere senso. Occorre limitarsi a definizioni finite, a considerare accanto ai razionali i reali «semplici», approssimabili a piacere, la cui totalità rimane comunque disperatamente numerabile cosicché il *continuo pratico* (come Borel chiama questo continuo che di fatto usiamo quando misuriamo grandezze fisiche) rimane una pallida approssimazione di quello geometrico. Borel oscilla tra la speranza che un giorno la matematica possa svilupparsi limitandosi a un continuo più dominabile e l'indagine di come si possano definire i singoli reali (classificando queste definizioni). È così che è condotto a considerare reali definiti empiricamente in termini di infinite scelte successive di cifre, scelte arbitrarie come quelle postulate dall'assioma di Zermelo.

Ma come è possibile che due matematici possano capire che si stanno riferendo allo stesso reale se esso è dato in termini di scelte infinite? L'intersoggettività verrebbe a cadere, non ci si capirebbe più e l'idea di successioni di libera scelta (o – peggio ancora – di scelte in quantità transfinita arbitraria, necessarie per dare un buon ordinamento al continuo) porterebbe al caos. Eppure – Borel lo ammette – solo accettando queste scelte il continuo pratico

potrebbe arricchirsi e colmare l'abisso che lo separa da quello geometrico; Borel non si stanca di sottolineare l'estrema complessità di questo concetto di numero reale, rimarcando ad esempio come mediante un'opportuna codifica sarebbe addirittura possibile tradurre tutto quanto mai scritto in un singolo numero reale, sicché in ultima istanza risulta impossibile parlare di *singoli* reali così intesi se non in senso probabilistico, considerando indiscernibili due reali che coincidono per un tratto del loro sviluppo decimale. Sino al 1917 è questa anche l'opinione di Brouwer, ma è proprio in quell'anno che egli individua una strategia che gli permetterà di conciliare la dicotomia di Borel e sviluppare una vera e propria teoria costruttiva del continuo che non si riduce a quello pratico (o «ridotto» come lo chiama Brouwer).

Per Brouwer, come per Borel, una successione di scelta è data da un'operazione che ad ogni naturale (e i naturali sono disponibili per un intuizionista) associa un oggetto appartenente ad una data *specie* (il concetto intuizionista che – come vedremo – corrisponde alla nozione classica di insieme). Il problema è: come è data questa operazione? Diversamente da Borel, Brouwer ammette che non è necessario che l'operazione in questione sia data da una legge (la «definizione finita» di Borel); successioni che sono così determinate si dicono anche *nomiche* (in inglese *lawlike*) ma non sono le uniche; esistono anche successioni *anomiche* (*lawless*) che ad esempio possono essere determinate utilizzando estrazioni casuali. Fra questi estremi si pongono successioni che vengono prodotte imponendo eventualmente vincoli sulle scelte future in termini di quelle passate, successioni cui si impongono vincoli sui vincoli, successioni *esitanti* in cui a un dato passo decidiamo che da lì in poi sceglieremo in base ad una legge oppure, se una decisione del genere non è già stata presa precedentemente, scegliamo liberamente, ecc. In lavori successivi Brouwer sottoporrà a perfezionamenti e revisioni le definizioni fondamentali, ma quel che rimane chiaro è che le successioni così individuate sono liberi atti creativi di un soggetto (il matematico ideale) che a partire da materiale già prodotto (ad esempio, i naturali nel loro svilupparsi) costruisce sistemi sempre più comprensivi senza vincoli esterni di adeguamento.

Se consideriamo successioni in cui l'operazione generatrice associa cifre da 0 a 9 a numeri naturali, oppure definizioni di successioni di Cauchy, otteniamo reali che costituiscono quel continuo

aritmetico (libero da vincoli geometrici) che Brouwer descrisse ripetutamente (ad esempio in *Begründung der Mengenlehre unabhängig von logischen Satz der ausgeschlossenen Dritten* [*Fondazione della teoria degli insiemi indipendente dall'assioma logico del terzo escluso*] del 1919 e nella conferenza di Vienna del 1928, poi pubblicata con il titolo *Die Struktur des Kontinuums* [*La struttura del continuo*]) sottolineandone la ricchezza rispetto a quello classico. Un segno di questa ricchezza è dato dalla validità per questo continuo pieno – ove si ammettono anche reali definiti da successioni arbitrarie di scelta – di principi molto forti, quale ad esempio la legge fondamentale per cui ogni funzione dal continuo C al continuo C è continua. Questo principio di continuità (noto come *principio di Brouwer*) non è valido classicamente in quanto, come fatto da Weierstrass, è possibile definire (classicamente) funzioni discontinue. Questo però non succede intuizionisticamente e Brouwer lo dimostra con la sua *tecnica dei controesempi*, facendo vedere che se esistesse una funzione discontinua definita sull'intervallo chiuso [0, 1] tale che $f(x) = 1$ se $x > 0$ e $f(0) = 0$ come quella definita da Weierstrass, potremmo ottenere una dimostrazione della congettura di Goldbach o della sua negazione. Questo è possibile appunto codificando il problema nella definizione di un opportuno reale nell'intervallo, addirittura fornendone una legge di definizione. La stessa sorte tocca al *teorema di Bolzano-Weierstrass*, al *principio di tricotomia* per i reali («per ogni reale x e y, $x > y$ o $x = y$ o $y > x$»), alla *proprietà di completezza* dei reali («se un insieme non vuoto di reali ha un confine superiore ha un minimo confine superiore») che classicamente caratterizza la continuità (come il lettore ricorderà dal cap. 3) eccetera.

Si noti che tutte queste proprietà classiche di R cessano di valere già nel continuo pratico o ridotto, i cui elementi sono dati da successioni nomiche, o da successioni di Cauchy, o da sviluppi decimali, ecc. Quello che l'aggiunta delle successioni di scelta fa è di rendere possibile *in positivo* la validità di principi classicamente non disponibili come il principio di continuità sopra richiamato. Come è possibile questo? L'idea di fondo di Brouwer è che se α è una successione di scelta per cui affermiamo $\mathscr{A}(\alpha)$, perché ciò abbia un senso dal punto di vista costruttivo occorre che a questa conclusione si giunga senza considerare α nella sua interezza, ma limitandosi a un suo *segmento finito* $\overline{\alpha(n)}$ (così indichiamo la sequen-

za $< \alpha(0), \alpha(1), \ldots, \alpha(n-1) >$, per $n \in \mathrm{N}$). Se quindi φ è una funzione che ad ogni reale del continuo esteso (quindi ad ogni successione di scelta) assegna un valore, è immediato concludere che

$$\forall\alpha\exists x\forall\beta\ (\overline{\beta(x)} = \overline{\alpha(x)} \rightarrow \varphi(\alpha) = \varphi(\beta)).$$

Questo non è altro che un modo di affermare che φ è continua una volta che il continuo N^{N} abbia la topologia di Baire (per la quale gli intorni di una α sono dati dalle famiglie $I(\alpha(n))$ di tutte le successioni β per cui $\overline{\beta(n)} = \overline{\alpha(n)}$ per ogni $n \in \mathrm{N}$).

Ritroviamo qui l'idea di Borel per cui non è possibile parlare in modo costruttivo di una *singola* successione di scelta, in quanto essa è indiscernibile dalle altre successioni con cui condivide un segmento iniziale. Il fatto nuovo – e rivoluzionario – è che Brouwer vede in ciò non l'impossibilità di considerare come oggetti matematici le successioni di scelta, ma la necessità di cambiare logica e di adottarne una adeguata a trattare oggetti «in fieri», oggetti incompleti quali le successioni che noi riusciamo ad afferrare e studiare solo attraverso i loro segmenti iniziali. È questa lettura degli enunciati matematici che permette di ottenere principi di continuità forti come quello sopra detto e di organizzare uno studio ricco e articolato del continuo su base costruttiva.

Quale sia questa nuova logica è presto detto: i principi che Heyting isola e giustifica nel 1933 in termini della sua lettura di connettivi e quantificatori erano stati saggiati precedentemente da Brouwer utilizzando proprio le successioni di scelta e il suo principio di continuità. Così, ad esempio, da quest'ultimo principio Brouwer può provare che

$$\neg\forall\alpha(\forall n\alpha(n) = 0 \vee \neg\forall n\alpha(n) = 0),$$

una vera e propria refutazione della validità universale del terzo escluso.

La matematica intuizionista non è quindi una semplice amputazione di quella classica (come credeva Hilbert) ma una matematica in cui le nozioni tradizionali si spezzano in concetti più fini, valgono principi non validi classicamente e principi classici valgono in forma modificata, acquistando quindi un altro significato. Ciò si verifica già per nozioni fondamentali e a prima vista

univoche come l'identità e l'ordine. Già abbiamo visto come nel continuo (ridotto o pieno) non valga in generale la tricotomia in quanto non esiste un metodo costruttivo in grado di confrontare rispetto all'ordine due qualsiasi successioni infinite (anche nomiche). A partire dal 1926 Brouwer sostituisce così al concetto di relazione d'ordine quello più generale di *ordine virtuale*. In generale, una specie S è virtualmente ordinata da una relazione $<$ asimmetrica se $<$ è un ordine parziale per cui valgono

$$\text{I)} \qquad x \nless y \wedge x \ngtr y \rightarrow x = y$$
$$\text{II)} \qquad x \nless y \wedge x \neq y \rightarrow x > y$$

Tanto I) che II) equivalgono classicamente alla tricotomia (cioè a $x < y \vee y < x \vee x = y$) ma così *non* è dal punto di vista intuizionista in cui $x \nless y$ sta per $\neg x < y$, $x \neq y$ per $\neg\, x = y$ che sono condizioni molto più forti della semplice non verità di $x < y$ e $x = y$. Così I) non ci permette di concludere che se $x < y$ e $y < x$ non sono vere allora $x = y$; la premessa di I) è più forte: esige che si sia dimostrato $\neg x < y$ e $\neg y < x$ e discorso analogo si può fare per II). Come Brouwer stabilisce nel 1927, l'ordine sul continuo è *virtuale* ed *inestendibile* nel senso che se $\neg a < b$ non è deducibile dagli assiomi d'ordine parziale, allora è deducibile $a < b$ e se non è deducibile $\neg a = b$ allora lo è $a = b$.

Questi fatti – come il venire a cadere della tricotomia – dipendono da come è definita $<$ sui reali. Trattandosi di successioni, affermare che $a < b$ significa fare riferimento al modo con cui sono date le due successioni, cioè al tipo di rappresentazione dei numeri reali di cui disponiamo. Brouwer analizzò in diverse occasioni la situazione rispetto alle sezioni di Dedekind, i decimali e le frazioni continue, anche se oggi forse il modo più diffuso di considerare i reali dal punto di vista costruttivo è quello in termini di successioni di Cauchy. Per ciascuna di queste rappresentazioni abbiamo le definizioni classiche della relazione $x < y$ in termini di elementi delle successioni, definizioni che però vanno lette costruttivamente. È quindi chiaro ad esempio che se α e β sono dati in termini di successioni di Cauchy e la relazione $<$ è definita nel solito modo, la verifica che $\alpha < \beta$ comporta l'individuazione effettiva di k, $n \in \mathrm{N}$ per cui $| \beta_{n+m} - \alpha_{n+m} | > 2^{-k}$ per ogni $m \in \mathrm{N}$. Non deve quindi stupire che non valga la tricotomia e che si impongano le

accurate distinzioni che Brouwer introduce. Queste distinzioni coinvolgono lo stesso concetto di identità che per i logicisti e i formalisti ha una portata così generale da poter essere assunto come concetto logico. Nell'interpretazione costruttiva questo non è vero perché le proprietà di $=$ cambiano a seconda degli oggetti che si considerano. Così, se ad esempio è ovvio che la relazione $=$ è decidibile tra naturali, interi o razionali, nel senso che possiamo decidere in questi casi se $a = b$ o $a \neq b$, ciò non accade se a, b, sono numeri reali. Anche qui dire che a e b sono lo stesso reale significa provare che (supponendo che essi siano dati mediante successioni di Cauchy) per ogni naturale k esiste un n tale che per ogni m

$$| \alpha_{n+m} - \beta_{n+m} | < 2^{-k}$$

Si vede così che in generale la negazione $\neg a = b$, ove $=$ non sia decidibile, è molto più forte del semplice fatto che non si riesce a provare $a = b$. L'identità tra reali gode di una proprietà più debole della decidibilità, cioè della *stabilità*

$$\neg\neg a = b \rightarrow a = b$$

mostrata già da Brouwer, e la sua negazione $\neq$ ha una controparte positiva, lo *scarto* (*apartness* in inglese) $\#$ definito ponendo

$$a \# b \text{ se e solo se } \exists k \exists n \forall m (\beta_{n+m} - \alpha_{n+m}) > 2^{-k}$$

e che ha le proprietà

$$\neg x \# y \rightarrow x = y$$
$$x \# y \rightarrow y \# x$$
$$x \# y \rightarrow x \# z \vee y \# z.$$

Si può provare che *ogni* relazione di scarto che soddisfa questi assiomi avrà associata un'identità stabile e la cosa (verificata da Brouwer per il caso dei reali) è stata studiata in generale da Heyting nel 1925, quando formulò tali assiomi e li usò come base prima del suo sviluppo dell'algebra lineare intuizionista nel 1927 e poi, a partire dal 1941, come punto di partenza per la costruzione di una teoria dei gruppi, degli anelli, dei campi, ecc., cioè di una vera e propria algebra intuizionista.

Già dagli anni venti Brouwer e Weyl avevano dato dimostra-

zioni costruttive del classico teorema fondamentale dell'algebra (secondo il quale i complessi sono un campo algebricamente chiuso) ma la cosa importante da osservare è che anche nel caso dell'algebra il tipo di costruttivismo che i concetti intuizionisti permettono di sviluppare è diverso da quello che risaliva a Kronecker e che sarà ripreso quasi simultaneamente in ambito hilbertiano da Grete Hermann e Van der Waerden (entrambi allievi della Noether). Secondo Kronecker l'algebra costruttiva deve limitarsi essenzialmente allo studio di strutture discrete su cui è definita un'identità decidibile e quindi era la stessa $\neq$ ad essere lo scarto fondamentale. Per l'intuizionismo, invece, il vincolo all'identità decidibile non è essenziale: quello che importa è ragionare con una logica adeguata a relazioni non decidibili come quella intuizionista, sicché $\neq$ è una relazione diversa da #. Si capisce, tenuto conto del carattere fondamentale di relazioni quali quella d'ordine e di identità nel definire la topologia sui reali, come l'Analisi intuizionista acquisti una maggiore sottigliezza di quella classica. Di alcuni teoremi classici che non valgono così come sono, si possono dare versioni intuizionisticamente accettabili che furono esplicitamente formulate e dimostrate da Brouwer stesso e dai suoi allievi (tra i quali M. Belifante, A. Heyting e J. Dijkman ecc.). Lo stesso vale per la teoria della misura, quella degli spazi metrici, la topologia generale e per la stessa topologia algebrica, cui Brouwer diede contributi fondamentali anche da un punto di vista non intuizionista.

Le nozioni generali sullo sfondo si spezzano e danno origine a distinzioni significative e sottili, prima di tutte quella d'insieme come collezione determinata da una proprietà che viene sostituita – come già accennato – da quella di *specie*. Dare una specie S significa dare una proprietà P per cui ha un *significato costruttivo* affermare che $P(a)$ per ogni a di una data collezione di oggetti A. È implicito che gli elementi di A sono dati *indipendentemente* da P, ed S sarà la collezione degli $a \in A$ per cui è possibile provare che $P(a)$. Come osservato a suo tempo, in questo caso non si pongono i problemi cui dava origine la nozione classica: A deve essere dato *prima* di S ed è implicita un'assunzione predicativista che ci impone di limitarci a proprietà P e a individui a per cui è *ben definita* costruttivamente la relazione «a gode di P». Di fatto, anche se sin dai suoi primi lavori Brouwer fece uso del concetto di specie, non

presentò mai un'analisi sistematica di quali proprietà possono loro dare origine. Sembra plausibile che questo sia il caso per le proprietà predicative, mentre la possibilità di usare anche proprietà non predicative è stata considerata solo successivamente. L'importante è che, anche prescindendo da problemi d'esistenza, ci sono basilari differenze tra l'approccio classico ed intuizionista alle collezioni.

Per cominciare, classicamente ogni insieme è determinato dai suoi elementi, vale a dire si assume un *principio di estensionalità*; questo non è accettabile dal punto di vista intuizionista, secondo il quale è essenziale *il modo* con cui un oggetto ci viene dato, sicché due specie (come due successioni di scelta) possono essere uguali estensionalmente ma non esserlo intensionalmente. Accanto alla *identità estensionale* abbiamo così quella *intensionale* o definizionale che comporta l'identità nel modo di presentare l'oggetto. Se quest'ultima è decidibile, la prima non lo è in generale; comunque l'importante è che la matematica intuizionista costituisce il primo esempio di una teoria in cui è essenziale distinguere tra proprietà, relazioni, ecc. che sono o meno estensionali. Ulteriori distinzioni sorgono quando consideriamo le proprietà basilari dell'appartenenza e dell'inclusione, anche queste dovute all'interpretazione costruttiva delle condizioni definitorie.

Così, se classicamente l'insieme X non è vuoto quando esiste un a per cui $a \in X$, costruttivamente la specie S sarà *abitata* solo se è possibile *costruire* un a per cui si può provare $a \in S$. Analogamente, mentre dal punto di vista classico X è *sottoinsieme* di Y ($X \subseteq Y$) se ogni elemento di X è elemento di Y, ora diremo che la specie X è *separabile in* Y se esiste una costruzione che per ogni $y \in Y$ prova che o $y \in X$ o $y \notin X$; in altri termini, si richiede, per poter considerare autonomamente X separandolo dall'ambiente Y, che l'appartenenza a X per ogni elemento di Y sia una proprietà *decidibile*.

Tutte queste distinzioni si ritrovano già nei lavori di Brouwer degli anni venti, ma solo col tempo, dopo che il processo di formalizzazione della matematica intuizionista sarà giunto a un notevole grado di sviluppo, possiamo parlare di una vera e propria teoria intuizionista delle specie, degli insiemi e delle costruzioni. L'attenzione maggiore di Brouwer e dell'intuizionismo negli anni trenta e quaranta è volta piuttosto ad altri concetti, in primo luogo a quello di *ordinale* e di *spiegamento* (inglese *spread*) che sono di-

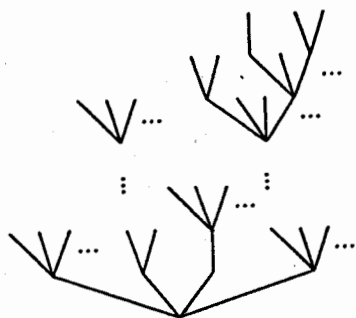
rettamente collegati all'analisi del continuo e al problema delle definizioni per induzione.

Va detto subito che per Brouwer, che non accetta in positivo la dimostrazione di Cantor per cui dato un insieme X il suo insieme delle parti (insieme potenza) ha cardinalità strettamente maggiore, esistono essenzialmente solo tre tipi di insiemi: quelli finiti, quelli numerabili come i naturali, e quelli con la cardinalità del continuo. Brouwer non fu sempre coerente in questa analisi e va inoltre aggiunto che il confronto cardinale tra insiemi dal punto di vista intuizionistico è estremamente più problematico che quello classico in quanto, almeno a partire dal 1925, il predicato di equipotenza si spezza in ben sette nozioni distinte, per cui cadono teoremi classici di confronto come quello di Schröder-Bernstein.

Quello che conta per quanto riguarda gli ordinali è che dal punto di vista intuizionista esistono solo ordinali numerabili, anche se la loro totalità non esaurisce la seconda classe numerica di Cantor, che per Brouwer non esiste. Dopo non poche esitazioni, nel 1926 Brouwer diede la definizione della specie *WO* delle specie bene ordinate che è tutt'oggi adottata. È una definizione induttiva di tipo genetico, come quella del primo Cantor: *WO* contiene la specie vuota (l'ordinale 0), è chiusa rispetto alla somma ordinale finita, al successore e se $\{F_n\}_{n\in N}$ è una successione *nomica* (quindi non arbitraria) di specie bene ordinate, anche la sua somma ordinale sarà in *WO*. Non c'è chiusura rispetto all'esponenziazione.

La definizione ci fornisce immediatamente un principio di induzione, ma l'aspetto più interessante della teoria emerge quando vediamo le specie bene ordinate in termini di successioni di naturali e facciamo ricorso al concetto di spiegamento che è la chiave di volta di tutta la teoria delle successioni e quindi tanto di quella degli ordinali quanto del continuo aritmetico. Consideriamo la specie S delle successioni *finite* di naturali che indicheremo con α, β, ecc.; possiamo definire un ordine $\leqslant$ sui suoi elementi ponendo che $\alpha \leqslant \beta$ se α è un segmento iniziale di β, cioè

$$\exists x(\alpha = \overline{\beta(x)}).$$

Così strutturata, la specie S diviene lo spiegamento universale o l'albero universale. È chiaro infatti che $\leqslant$ induce una struttura ad albero

la cui *radice* è la successione vuota $< >$ e per ogni vertice α i successori sono le successioni che si ottengono concatenando $\alpha = < n_1, ..., n_k >$ con n, ottenendo $\alpha * n = <n_1, ..., n_k, n>$ per ogni $n \in$ N. Rami dell'albero sono tutte quelle successioni α le cui sequenze iniziali $\overline{\alpha(n)}$ appartengono tutte a S. È chiaro che ogni successione $\alpha \in N^N$ sarà un ramo dell'albero universale e l'interesse di vederle come elementi dello spiegamento universale è che abbiamo un modo «visivo» di seguire il processo di crescita delle successioni studiandone le alternative di sviluppo ad ogni passo.

Il concetto generale di spiegamento diviene significativo una volta che vogliamo considerare successioni o segmenti che appartengono a una data sottospecie di S. Diremo allora che T è un albero se è una sottospecie abitata, separabile di S (che costituisce a sua volta un albero) che contiene la successione vuota e che se contiene una successione contiene tutti i suoi segmenti iniziali. Parleremo di *spiegamento* se T è un albero in cui ogni vertice ha almeno un successore, cioè per ogni $\alpha \in T$ esiste almeno un $n \in$ N per cui $\alpha * n \in T$. Si noti che un vertice può avere più di un successore; in S, ad esempio, ogni vertice ha infiniti successori.

Diremo che T è un *ventaglio* (o spiegamento finitario) se ha ramificazioni finite, cioè se ogni vertice ha un numero *finito* di successori. Un ventaglio quindi, come ogni spiegamento, ha infiniti vertici ma ogni vertice ha solo un numero finito di successori. Come si ricorderà, per strutture di questo tipo esiste un teorema classico dimostrato da König nel 1927 che afferma che in un albero finitario (un ventaglio) esiste almeno un ramo infinito. La dimostrazione fa uso di una forma debole dell'assioma di scelta e consiste essenzialmente nel definire per induzione su N una successione che ad ogni passo a partire dalla radice dell'albero sceglie un suc-

cessore del vertice scelto precedentemente. Come tale il teorema è uno strumento essenziale (dal punto di vista classico) per costruire oggetti (insiemi, funzioni, strutture, ecc.) definibili per induzione su N in situazioni in cui a ogni passo i prolungamenti costituiscano una diramazione finita. È proprio un principio come questo che – abbiamo visto – Skolem ed Herbrand si rifiutarono di utilizzare per globalizzare le soluzioni date livello per livello e cui Gödel consapevolmente fece riferimento per costruire contromodelli nella sua dimostrazione di completezza per la logica del primo ordine. Skolem ed Herbrand non fecero il passo per scrupoli di carattere finitario (o intuizionista, come pensava Herbrand). La domanda che si pone è: c'è un sostituto intuizionisticamente corretto per il teorema di König? A che principi rimanda e che rapporti ha con l'assioma di scelta?

Il problema ha un'importanza fondamentale in quanto è strettamente connesso alla possibilità di costruire successioni di scelta che rispondono a date caratteristiche e più in generale al significato intuizionista delle procedure induttive.

Nel 1927 Brouwer risolse la questione dimostrando un equivalente classico dell'*Unendlichkeit lemma* di König, il cosiddetto *teorema del ventaglio*, che è una sorta di versione contrapposta di quello di König. Il teorema afferma che se abbiamo un ventaglio V e una sottospecie separabile R che forma uno *sbarramento* (nel senso che ogni ramo del ventaglio che parte dall'origine viene ad un certo punto a passare per R) allora esiste un naturale n tale che tutti i rami incontrano un punto di R dopo al massimo n passi. Se consideriamo alberi di sequenze finite, identifichiamo i cammini finiti del lemma di König con i rami che da un certo punto in poi contengono solo sequenze finite la cui parte finale ha solo zeri e consideriamo come R la specie delle sequenze di questo tipo, è ovvio che il teorema del ventaglio è letteralmente il contrapposto di quello di König; che il secondo non si possa ottenere intuizionisticamente dal primo è chiaro proprio perché il principio di contrapposizione non vale intuizionisticamente.

Per dimostrare il teorema Brouwer introdusse un principio che gli permetteva di procedere per induzione sul ventaglio e che da allora è divenuto il pilastro di tutta la teoria intuizionista delle definizioni induttive generalizzate (e non solo di quella, come vedremo, considerando i lavori di C. Spector). Il principio si riferisce a

spiegamenti arbitrari S e non necessariamente a ventagli. Esso afferma che se S è uno spiegamento, R una sottospecie separabile che sbarra S, A una sottospecie di S che estende R per cui

$$\forall u \in S(\forall n \in \mathrm{N}\ u * n \in A \rightarrow u \in A)$$

allora la radice $< >$ appartiene ad A, $< > \in A$.

Assunto questo principio, la dimostrazione del teorema del ventaglio è quasi immediata se si prende come A la specie dei livelli u per cui tutti i rami che passano per u si trovano in R dopo al massimo n passi per un prefissato $n \in \mathrm{N}$. Brouwer cercò più volte di giustificare dal punto di vista costruttivo il principio e da allora tanto esso che la dimostrazione del teorema del ventaglio sono stati oggetto di analisi e di discussioni (in tempi recenti, ad esempio, da parte di M. Dummett).

Quello che a noi interessa sottolineare è che il teorema del ventaglio costituisce la chiave di volta dell'Analisi e della teoria del continuo intuizioniste come hanno mostrato lo stesso Brouwer e, più recentemente, i lavori ad esempio di A. Troelstra, D. van Dalen, van Rootselaar, e altri. La conseguenza più sorprendente è forse costituita dal teorema per cui ogni funzione definita su un intervallo chiuso (del continuo pieno) è *uniformemente* continua. Questo è un rafforzamento del principio di continuità di Brouwer, non valido classicamente, che mostra la ricchezza dell'Analisi intuizionista ed è un esempio di quei principi di continuità che discendono combinando interpretazione costruttivista, sequenze di scelta, teorema del ventaglio e assioma di scelta. Perché, se l'assioma è chiaramente inaccettabile dal punto di vista costruttivo se applicato ad insiemi qualsiasi, non lo è più se costruttivamente interpretato. È sostanzialmente in un lavoro di Kleene del 1957 che per la prima volta si assume una forma dell'assioma come base di una teoria intuizionista, ma l'analisi più significativa dell'assioma in questo contesto si trova in alcuni fondamentali lavori di Kreisel del 1963 sulle successioni di scelta e di Kreisel e Howard, e Kreisel e Troelstra del 1970.

In anni più tardi, a partire dal 1948, Brouwer introdusse altri principi basati su nozioni più sfuggenti quali quella di *soggetto creativo* che ha dato origine a numerose discussioni ed elaborazioni. L'idea è – ancora una volta – quella di considerare il conoscere nel suo sviluppo e così la relazione di fondo è quella che – simboleg-

giata con $\vdash_m \mathscr{A}$ – sussiste quando il soggetto allo stadio m ha una prova per $\mathscr{A}$. Che principi valgono per questa relazione? Le alternative sono diverse e comportano conseguenze anche sul piano dell'Analisi intuizionista una volta che (come proposto da Myhill nel 1968, sulla falsariga del metodo brouweriano dei controesempi) si definisce *matematica* una successione nomica che non coinvolge $\vdash_m$ nella sua definizione.

La teoria del soggetto creativo ha trovato interessanti collegamenti con altri principi che partono dalla considerazione dello sviluppo della conoscenza. Così è ad esempio per il principio proposto da Markov nel 1954, per cui

$$\forall x(\mathscr{A}x) \vee \neg\mathscr{A}(x)) \rightarrow \neg\neg\exists x\mathscr{A}(x) \rightarrow \exists x\mathscr{A}(x)$$

vale a dire (interpretando costruttivamente) che se $\mathscr{A}$ è una proprietà decidibile e si ha prova che porta ad assurdo l'ipotesi $\neg\exists x\mathscr{A}(x)$, allora $\exists x\mathscr{A}(x)$. Il principio è plausibile una volta che – come faceva Markov – lo si legga in termini di algoritmi: se è impossibile che un algoritmo non termini la computazione per un dato input (cioè se è impossibile che non *esista* un ultimo passo della computazione) allora la computazione stessa termina, esiste un ultimo passo. La cosa è ovvia in termini di macchine di Turing, il concetto di macchina calcolatrice astratta che – come vedremo – fu proposto negli anni trenta come precisazione della nozione di computabilità; ma lo è nel contesto intuizionista?

Un argomento di Brouwer risalente al 1948 sembra provare che il principio di Markov è incompatibile intuizionisticamente con la teoria del soggetto creativo una volta che si assuma la validità del cosiddetto *principio di Kripke* secondo il quale

$$\forall X\, \exists\alpha\forall n\; (n \in X \rightarrow \exists m\alpha(n * m) = 0)$$

e che fu presentato per la prima volta da Myhill nel 1967.

Questo è la spia di una situazione più generale, di cui ci dà ulteriore conferma il fatto che – se ci limitiamo a successioni date ricorsivamente – il teorema del ventaglio è falso, come è stato dimostrato da Kleene.

L'intuizionismo non è l'unico modo di sviluppare una matematica costruttiva. Altri ne esistono – che possiamo solo ricordare – che si sono affermati prima dell'intervento di Brouwer, come il

predicativismo di Poincaré e Weyl, o dopo di esso. Così è per la *matematica ricorsiva* propugnata da Markov (1903-79) a partire dal 1950 che assume il principio di Markov di sopra, utilizza una sua interpretazione dei connettivi e dei quantificatori e soprattutto considera come oggetti – sulla scia di Kronecker – solo quelli codificabili con algoritmi, questi ultimi identificati con macchine di Turing (di cui parleremo più avanti). Così è pure per il costruttivismo di Bishop il cui manifesto è il libro *Foundations of constructive Analysis* (*Fondamenti di Analisi costruttiva*) pubblicato nel 1968. Qui l'idea di fondo è che ogni enunciato matematico deve avere significato numerico, ma non si identificano gli oggetti con quelli dati algoritmicamente; vengono accettati concetti astratti come quello di regola o costruzione, ma non le successioni di scelta di Brouwer, mentre si utilizza un'interpretazione costruttiva degli operatori logici. Su questa base Bishop riesce a dare una versione costruttiva di gran parte della teoria della misura, dell'Analisi funzionale, ecc. Esistono infine tentativi di matematica costruttiva che si muovono all'*interno* della matematica classica, lungo la via di Kronecker. Risultati interessanti in questo senso sono stati ottenuti in tempi recenti da A. Nerode e G. Metakides, F. Richman, ecc., il cui obiettivo è di isolare sistematicamente il significato *costruttivo* di teoremi classici effettivizzando le nozioni tradizionali (i primi lavori in questo senso per quanto riguarda l'algebra sono di M. Rabin nel 1960 e sono collegati ai classici problemi di decisione che sorgono nella teoria dei gruppi). Non deve stupire quindi che W. Tait abbia sostenuto nel 1985 che la matematica costruttiva è parte di quella classica e non coincide affatto con l'intuizionista.

Non tutti questi sviluppi coinvolgono direttamente la logica ma pongono sicuramente problemi di carattere logico che divennero affrontabili dopo i primi lavori di formalizzazione della matematica intuizionista ad opera di Heyting. Come confrontare la forza della matematica classica e dei vari indirizzi costruttivi? Quali principi costruttivi sono compatibili? A queste domande poteva dare risposta solo un'indagine metamatematica e a partire dagli anni sessanta si sono avuti sviluppi estremamente interessanti al riguardo che – come vedremo più avanti – hanno coinvolto teoria della dimostrazione, ricorsività, informatica e teoria delle categorie, permettendo di ampliare lo stesso orizzonte costruttivista e di stabilire più profondi legami con l'indagine classica.

5. LA «REVISIONE» DEL PROGRAMMA HILBERTIANO. GERHARD GENTZEN

Se nello sviluppo dell'intuizionismo la formalizzazione ebbe un effetto che solo col tempo si andò consolidando, già alla fine degli anni trenta la possibilità di confrontare direttamente le due concezioni che prendevano sul serio gli aspetti costruttivi – formalismo e intuizionismo – e la comparsa dei risultati di Gödel portò alcuni logici e matematici a ripensare *ex novo* il problema della costruttività e del finitismo. L'urgenza di un ripensamento era tanto più chiara dopo i risultati di Gödel sull'incompletezza e indecidibilità dell'aritmetica. È su questo sfondo che si situano gli interventi già ricordati di H. Weyl e soprattutto quelli di Gerhard Gentzen (1909-45) per noi in questo contesto più decisivi in quanto portarono ad un nuovo modo di concepire la teoria della dimostrazione e il suo ruolo nell'indagine sui fondamenti.

Gentzen riaffronta il problema dei fondamenti della matematica a partire da quella che – hilbertianamente – considera la sua radice, ossia dalle antinomie. Non si tratta di analizzarle in termini di «qualcosa che non va» nel nostro pensiero: questo atteggiamento comporta in modo naturale tutte quelle proposte di superamento fondate sul tentativo di tracciare una linea di demarcazione fra forme di inferenza matematica ammissibili e non ammissibili. Questi tentativi si sono senz'altro dimostrati utili nella pratica – sostiene Gentzen – ma non sono adeguati da un punto di vista teorico: da un lato infatti non si sa con esattezza *quale* sia l'eventuale errore, d'altro lato, una volta operata la distinzione precedente, il problema si ripropone, talora in forma addirittura più acuta. «La situazione», dice Gentzen, «è piuttosto che risulta impossibile parlare di un errore univocamente identificabile nel nostro pensiero. Tutto ciò che può essere detto con certezza è che l'apparizione delle antinomie è connessa al concetto di infinito».

Riprendendo quindi una classificazione di Hilbert (che a sua volta si era ispirato a Weyl) egli vede la matematica suddivisa in tre livelli, a seconda del (l'estensione del) concetto di infinito che in essa viene accettato: avremo così un primo livello costituito dalla teoria elementare dei numeri (ossia l'aritmetica, la teoria dei numeri naturali), un secondo livello costituito dall'Analisi (ossia dalla teoria dei numeri reali, vale a dire dalla teoria di insiemi

qualunque di numeri naturali) e infine un terzo livello costituito dalla teoria generale degli insiemi (ossia dall'intera teoria dei numeri cardinali e ordinali di Cantor). Nel motivare come detto il presentarsi delle antinomie, Gentzen concorda evidentemente con i massimi esponenti dell'intuizionismo in senso lato, Poincaré, Weyl, Brouwer ecc., per i quali la cosa era sostanzialmente riconducibile, come sappiamo, alla stessa concezione *attualistica* (come Gentzen la chiama) dell'infinito; non ritiene tuttavia di poter aderire *in toto* al loro tipo di costruttivismo radicale quando essi dichiarano semplicemente *sprovvista di senso ogni* proposizione matematica nella quale si abbiano riferimenti «infinitari» (nella quale cioè vengano quantificate, universalmente o esistenzialmente, variabili che hanno un dominio di variabilità infinito). Se si dovesse adottare questo punto di vista, sostiene Gentzen, «l'intera Analisi classica si ridurrebbe a un campo di macerie. Molti teoremi fondamentali perderebbero la loro validità o dovrebbero essere riformulati e ridimostrati in modo diverso. Senza contare che la formulazione diverrebbe nella maggior parte dei casi *più complicata* e le dimostrazioni *più noiose*».

Si tratta piuttosto, secondo Gentzen, in adesione con la concezione hilbertiana, di dare un *significato finitista* (e quindi a fortiori *costruttivo*) a proposizioni di questo tipo, interpretando in particolare in modo costruttivo, vale a dire facendo riferimento al solo infinito *potenziale*, quelli che sono i riferimenti infiniti *attualisti* compresi in quelle proposizioni. Ovviamente, nella radicale posizione intuizionista, anche una dimostrazione di coerenza perde completamente di significato. Come sappiamo, invece, dal punto di vista di Hilbert è proprio questo il passo che giustifica l'intero procedimento, che permette cioè di operare con «proposizioni ideali» *come se* esse esprimessero un contenuto e avessero quindi un senso. Tuttavia, il punto di vista strettamente finitista nel senso di Hilbert non è più sostenibile dopo la dimostrazione dei teoremi di Gödel; in particolare per quanto riguarda le dimostrazioni di coerenza, osserva Gentzen «sembra siano richiesti metodi alquanto più forti di quelli originariamente accettati da Hilbert e che egli aveva pensato costituire le "tecniche finitiste di dimostrazione"». Occorre in altri termini in qualche modo «rafforzare» il finitismo, senza tuttavia uscire dall'ambito del costruttivo; occorrerà far vedere che i nuovi metodi, cui si deve necessariamente ricor-

rere dopo il teorema di Gödel, rimangono, pur se più potenti, sempre «in armonia con l'*interpretazione costruttivistica* dell'infinito».

Il punto centrale di tutto il discorso rimane quindi, per Gentzen, sostanzialmente lo stesso al quale era giunto Hilbert: occorre far diventare la stessa *dimostrazione matematica* un oggetto matematico suscettibile di indagine rigorosa, formale. Ed è dalla concezione stessa di dimostrazione che è necessario far ripartire l'analisi. Non è quindi casuale che già nel 1932 in *Über die Existenz unabhängiger Axiomensysteme zu unendlichen Satzsystemen* (*Sull'esistenza di sistemi di assiomi indipendenti per sistemi infiniti di proposizioni*), riprendendo alcune idee che il logico e matematico Paul Hertz aveva avanzato in una serie di lavori negli anni 1928-1929, Gentzen prenda in esame il concetto stesso di dimostrazione informale, tentando di analizzarlo, a livello formale, in modo più fedele a quella che era l'usuale pratica del ragionamento matematico; come è assai significativo che nello stesso lavoro, accanto agli specifici risultati raggiunti,[19] egli intraprenda (anteriormente quindi allo stesso Tarski) una prima caratterizzazione della nozione di *conseguenza logica* che avviene appunto sulla base di un nuovo modo di analizzare il processo dimostrativo informale ed è particolarmente interessante in quanto fornisce, almeno in linea di principio, un'analisi della nozione semantica di conseguenza sulla base di considerazioni *sintattiche e «finitarie»*.

Ma è nel 1935, in *Untersuchungen über das logischen Schliessen* (*Ricerche sulla deduzione logica*), che Gentzen approfondisce e articola il suo discorso in modo del tutto soddisfacente.[20] Anche qui egli vuol

[19] Gentzen risponde in modo definitivo e completo a un serio problema allora aperto e indicato dal titolo stesso del suo lavoro: l'esistenza di assiomatizzazioni indipendenti per sistemi arbitrari di proposizioni. Egli mostra con un controesempio che non tutti questi sistemi ammettono assiomatizzazioni di questo tipo, mentre individua i sistemi infiniti per i quali ciò avviene nei sistemi costituiti da quelle che lui chiama proposizioni «lineari», che caratterizza in modo completo.

[20] È qui opportuno ricordare che un'impostazione analoga a quella di Gentzen aveva portato il polacco Stanislaw Jaśkowski a costruire una «teoria della deduzione basata sul metodo delle supposizioni», che aveva pubblicato nel 1934 in *On the rules of suppositions in formal logic* (*Sulle regole delle supposizioni in logica formale*), ma la cui idea, che Jaśkowski accredita a Łukasiewicz, risaliva almeno alla metà degli anni venti. La teoria che ne risulta è tuttavia alquanto farraginosa e assai meno trasparente di quella di Gentzen, che può senz'altro ritenersi indipendente. È anche interessante osservare – per ribadire la caratteristica tendenza della ricerca logica in Gentzen – che questi, nel 1933, quando aveva già corretto le bozze dell'articolo *Über das Verhältnis zwischen intuitionistischer und klassischer Arithmetik*

trattare tanto la logica classica che quella intuizionista, e precisamente con quella parte della logica, il calcolo dei predicati del primo ordine, «che comprende i tipi di inferenza continuamente impiegati in tutta la matematica»; solo che, come dicevamo, egli tende a una formalizzazione di questo calcolo che risulti più aderente all'effettiva pratica del ragionamento matematico: «Il mio punto di partenza è stato questo: la formalizzazione della deduzione logica, in particolare come è stata sviluppata da Frege, Russell e Hilbert, si discosta alquanto dalle forme di deduzione usate nella pratica matematica delle dimostrazioni. In cambio se ne ottengono considerevoli vantaggi formali. Al contrario era mia intenzione prima di tutto costituire un sistema formale che si ponesse quanto più possibile vicino all'effettivo ragionamento. Il risultato è stato un "*calcolo della deduzione naturale*"... che viene ad avere certe proprietà peculiari; in particolare, la "legge del terzo escluso" che gli intuizionisti rifiutano, occupa una posizione speciale.»

L'idea di fondo dell'analisi delle dimostrazioni informali che Gentzen intraprende è che esse in generale prendono le mosse non tanto da poche formule logiche fondamentali (gli assiomi) e si sviluppano applicando poche regole di inferenza, quanto, esattamente al contrario, impiegando se occorre numerose *assunzioni* necessarie alla dimostrazione della proposizione cui si è interessati, e

(*Sul rapporto fra l'aritmetica intuizionista e quella classica*) lo ritira dalla pubblicazione (sui *Mathematische Annalen*) non appena conosciuto l'analogo articolo di Gödel uscito poco prima, e che noi abbiamo visto nel paragrafo 4. In questo lavoro Gentzen era giunto sostanzialmente agli stessi risultati di Gödel, vale a dire che l'aritmetica classica differisce da quella intuizionista in modo puramente esteriore (in pratica: solo perché è fondata su un *diverso* linguaggio logico predicativo). È ovvio, già da quanto abbiamo sopra detto, che per Gentzen si imponesse l'utilità – se non la necessità – delle ricerche su questi rapporti, che del resto saranno una componente si può dire costante di tutta la sua produzione. È anche interessante leggere, alla luce di questo fatto, le sue conclusioni in questo lavoro: «Se si accetta come coerente l'aritmetica intuizionista, allora è anche garantita la coerenza dell'aritmetica classica... Se si assume come punto di partenza un punto di vista più ristretto, come ad esempio quello "finitista" di Hilbert da lui delineato nell'articolo "Sull'infinito", allora ci resta anche il compito di dimostrare, da questo punto di vista, la coerenza dell'aritmetica intuizionista. È alquanto dubbio che questo possa farsi, perché Gödel ha dimostrato che la coerenza dell'aritmetica classica (più precisamente: una proposizione aritmetica equivalente) non può essere dimostrata nell'aritmetica stessa (ammesso che l'aritmetica sia coerente)... Applicando il risultato di Gödel [del 1931] concludiamo che la coerenza dell'aritmetica *intuizionista* non può essere dimostrata nell'aritmetica *classica* (assumendola coerente).»

operando su tali assunzioni con tutta una serie di «regole» che sostanzialmente non fanno altro che collegare («connettere») proposizioni già assunte o dimostrate, operare in esse sostituzioni, affermarne la validità in generale, passare da un esempio a una proposizione esistenziale e così via; applicando ovviamente, se necessario, anche le operazioni inverse. In altri termini, si tratta di indagare gli operatori logici nel loro stesso modo di essere *impiegati* informalmente, per poterne poi trarre una sistemazione formale «naturale». E se si conduce – come Gentzen fa – un'analisi di dimostrazioni matematiche concrete, ci si accorge che l'*uso* che in esse viene fatto degli operatori logici riguarda sostanzialmente due tipi di «manipolazioni» di fondo: la loro introduzione e la loro eliminazione. Nel primo caso infatti ad esempio, ottenute sulla base di certe assunzioni due proposizioni $\mathscr{A}$ e $\mathscr{B}$ separatamente, si passerà a connetterle variamente fra loro per ottenere una sola proposizione complessa; il secondo caso rappresenta l'operazione esattamente inversa. Ne viene che per formalizzare adeguatamente e naturalmente il processo dimostrativo occorrerà regolare opportunamente proprio questo comportamento degli operatori logici.

Si noti che così facendo Gentzen riduce la dimostrazione a un oggetto matematico più strutturato di quanto accadeva con le precisazioni di Frege, Russell e Hilbert (in particolare, ovviamente, di quest'ultimo); ora infatti una dimostrazione diviene sostanzialmente un *albero* le cui «radici» non sono più delle formule particolari e fisse (gli assiomi) ma delle formule qualsiasi (le assunzioni) e il cui andamento non è lineare e rigidamente fissato da poche regole estremamente artificiose rispetto all'effettivo andamento della dimostrazione intuitiva stessa, bensì segue quest'ultima passo passo, evidenziando per così dire con varie diramazioni (i «rami») le dipendenze «reali», fra assunzioni e conclusione.[21] Non si può certamente pensare a un'analisi più «costruttiva» in senso intuizionista; e in effetti la questione che si potrebbe qui sollevare è proprio la seguente: è stata l'esigenza di tenere nel dovuto conto la concezione intuizionista che ha condotto Gentzen a questo tipo di

[21] Va detto tuttavia che in alcune sistemazioni odierne dei calcoli di Gentzen si preferisce in generale dare una disposizione *lineare* alle derivazioni; noi ci riferiremo qui a questo tipo di sistemazione. Si veda in proposito nella successiva nota un esempio di derivazione di **NJ**.

analisi, o viceversa l'esigenza di maggior adeguamento alla corrente pratica matematica ha dato (e non casualmente) origine a un'analisi tanto «naturale» e vicina alla concezione intuizionista? Lasciamo comunque da parte questo interessante problema e vediamo brevemente come Gentzen realizza i suoi calcoli.

Nelle *Ricerche* Gentzen introduce due tipi di calcoli: i calcoli **N** (**NJ** per il calcolo intuizionista, **NK** per il calcolo classico) e i calcoli **L** (ancora **LJ** e **LK**). È per i primi, in particolare, che viene realizzata quella «naturalezza» di cui si è parlato, mentre i secondi, ossia i calcoli **L**, a prezzo di qualche complicazione, gli permettono di ottenere risultati ancor oggi fondamentali nella teoria della dimostrazione. In entrambi i casi si tratta di calcoli predicativi del primo ordine. Per quanto riguarda **NJ** si avranno per i cinque connettivi e i due quantificatori 14 regole, due per ogni operatore, una di introduzione e una di eliminazione. Esemplifichiamo tali regole per il connettivo di implicazione. Se $\mathscr{A}_i$, $\mathscr{B}$, $\mathscr{C}$, sono formule, le regole di introduzione ($I \rightarrow$) e di eliminazione ($E \rightarrow$) sono:

$$I \rightarrow \frac{[\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_n, \mathscr{B}] \quad \mathscr{C}}{[\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_n]\, \mathscr{B} \rightarrow \mathscr{C}} \qquad\qquad E \rightarrow \frac{\begin{array}{ll}[\mathscr{A}_1, \ldots, \mathscr{A}_m] & \mathscr{B} \rightarrow \mathscr{C} \\ [\mathscr{A}_{m+1}, \ldots, \mathscr{A}_n] & \mathscr{B}\end{array}}{[\mathscr{A}_1, \ldots, \mathscr{A}_n] \quad \mathscr{C}}$$

da leggersi, rispettivamente: se (nel corso di una derivazione) sotto le assunzioni $\mathscr{A}_1, \ldots, \mathscr{A}_n$, $\mathscr{B}$ si è ottenuta la formula $\mathscr{C}$ (che ovviamente può coincidere con una delle $\mathscr{A}_i$ o con $\mathscr{B}$), allora sotto le assunzioni $\mathscr{A}_1, \ldots, \mathscr{A}_n$ si può ottenere la formula $\mathscr{B} \rightarrow \mathscr{C}$, vale a dire si può *scaricare* una assunzione $\mathscr{B}$ fatta nel corso di una derivazione ponendola come antecedente di un condizionale il cui conseguente sia la formula ($\mathscr{C}$) precedentemente ottenuta; naturalmente si può applicare questa regola un numero qualsiasi finito di volte fino a «scaricare» *tutte* le assunzioni fatte per ottenere una certa conclusione (il che noi indicheremo, all'occorrenza, con [—]). Il contenuto di questa regola è l'equivalente del cosiddetto *teorema di deduzione* valido (sotto opportune condizioni) per la sistemazione assiomatica della logica di tipo hilbertiano. Per quanto riguarda $E \rightarrow$ si vede immediatamente che essa corrisponde al *modus ponens* e la sua lettura non dovrebbe presentare problemi. Dice Gentzen: «Dal punto di vista esteriore, la differenza essenziale fra le dimo-

strazioni in **NJ** e quelle nei sistemi di Russell, Hilbert e Heyting è la seguente: in questi sistemi le formule vere sono derivate da un insieme di "formule logiche fondamentali" mediante poche forme di inferenza. La deduzione naturale invece non parte in generale da proposizioni logiche fondamentali ma piuttosto *da assunzioni*... alle quali vengono applicate le deduzioni logiche. Per mezzo di successive inferenze il risultato viene quindi reso indipendente dalle assunzioni».[22]

Così formulato il calcolo, è possibile individuare ruolo e carattere del tutto speciale del terzo escluso $\mathscr{A} \vee \neg \mathscr{A}$: si passa dal calcolo intuizionista **NJ** a quello classico **NK** proprio ammettendo, in quest'ultimo, di poter assumere come *assioma* (ossia sotto nessuna assunzione) in qualunque «riga» di qualunque derivazione, appunto la formula $\mathscr{A} \vee \neg \mathscr{A}$; o, se si preferisce, ciò equivale ad aggiungere agli schemi di inferenza per gli operatori logici un ulteriore schema, che permette l'eliminazione della doppia negazione ed è così esprimibile sinteticamente: $\neg \neg \mathscr{A} / \mathscr{A}$ (come sappiamo questo passaggio non è intuizionisticamente valido, ed è appunto equivalente al terzo escluso).

Più duttili dal punto di vista metateorico, ma meno intuitivi sono i calcoli **LJ** e **LK**, in cui Gentzen non si riferisce semplicemente

[22] Diamo qui un esempio di derivazione in **NJ**, dimostrando la formula proposizionale $(q \rightarrow r) \rightarrow ((p \vee q) \rightarrow (p \vee r))$. Con i concisi chiarimenti che faremo seguire il lettore dovrebbe essere in grado di «ricostruire» le regole applicate.

1.	$[q \rightarrow r]$	$q \rightarrow r$	Ass.
2.	$[q]$	q	Ass.
3.	$[q \rightarrow r, q]$	r	$E \rightarrow$, 1, 2
4.	$[q \rightarrow r, q]$	$p \vee r$	$I \vee$, 3
5.	$[p]$	p	Ass.
6.	$[p]$	$p \vee r$	$I \vee$, 5
7.	$[p \vee q]$	$p \vee q$	Ass.
8.	$[q \rightarrow r, p \vee q]$	$p \vee r$	$E \vee$, 4, 6, 7
9.	$[q \rightarrow r]$	$(p \vee q) \rightarrow (p \vee r)$	$I \rightarrow$, 8
10.	$[-]$	$(q \rightarrow r) \rightarrow$ $\rightarrow ((p \vee q) \rightarrow (p \vee r))$	$I \rightarrow$, 9

Nelle righe 1, 2, 5 sono state introdotte delle assunzioni, il che è indicato con un Ass. a destra delle righe corrispondenti. Nella riga 3, ad esempio, è stata applicata la regola di eliminazione di $\rightarrow$ alle (formule delle) righe 1, 2. La «meno ricostruibile» delle regole sopra applicate è forse la $E\vee$. Essa si basa su questo elementare ragionamento: se entrambi i corni p, q di un dilemma ($p \vee q$, riga 7) permettono di derivare (eventualmente con altre assunzioni) una stessa formula (nel nostro caso, $p \vee r$, righe 6, rispettivamente, 4) allora possiamo eliminare il dilemma stesso affermando la formula così ottenuta (riga 8).

a formule, ma considera il concetto, fondamentale in questo contesto, di *sequenza* di formule, che può indicarsi ad esempio con

$$(1) \qquad \mathscr{A}_1, ..., \mathscr{A}_k \vdash \mathscr{B}_1, ..., \mathscr{B}_m \qquad (k \geqslant 0, m \geqslant 0)$$

dove le $\mathscr{A}_i$ (*antecedente* della sequenza) e le $\mathscr{B}_i$ (*conseguente*) sono *formule*; una sequenza come la (1) va letta: almeno una delle formule $\mathscr{B}_i$ segue, è derivabile, dalle formule $\mathscr{A}_i$, ossia, intuitivamente, va intesa come equisignificante con la formula $(\mathscr{A}_1 \wedge \mathscr{A}_2 \wedge ... \wedge \mathscr{A}_k) \rightarrow (\mathscr{B}_1 \vee \mathscr{B}_2 \vee ... \vee \mathscr{B}_m)$. Possono aversi sequenze con antecedente vuoto ($k = 0$) che si riducono alla formula $\mathscr{B}_1 \vee \mathscr{B}_2 \vee ... \vee \mathscr{B}_m$ (ossia: vale almeno una delle formule $\mathscr{B}_i$); o sequenze con conseguente vuoto ($m = 0$) che si riducono alla formula $\neg (\mathscr{A}_1 \wedge \mathscr{A}_2 \wedge ... \wedge \mathscr{A}_k)$, vale a dire l'antecedente vuoto significa una assunzione vera, il conseguente vuoto una conclusione falsa. Può anche aversi una sequenza vuota che scriveremo ad esempio «$\vdash$» e che è equivalente a una contraddizione (da una premessa vera si è ottenuta una conclusione falsa). Gli unici «assiomi» ammessi da Gentzen sono sequenze del tipo $\mathscr{A} \vdash \mathscr{A}$.

La differenza fra i calcoli **LJ** e **LK** sta ora in questo che, restando comuni a entrambi gli schemi inferenziali cui accenneremo, nelle derivazioni in **LJ** sono ammesse *solo* sequenze il cui conseguente sia costituito da *una sola formula*. Aggiungendo tale limitazione, e salvo esplicito avviso in contrario, quanto ora diremo genericamente sui calcoli **L** varrà tanto per **LJ** quanto per **LK**.

Per stabilire un calcolo **L** delle sequenze (o, come anche si dice, dei *sequenti*), si dovranno dare, come per i calcoli **N**, delle regole per l'introduzione e l'eliminazione dei connettivi logici (il che si può ridurre, come ci si convince facilmente, alla sola *introduzione* dell'operatore in questione nell'antecedente *o* nel conseguente di una sequenza); ma saranno anche necessarie regole che riguardano soltanto la «struttura» di una sequenza, nel senso che permettono ad esempio di scambiare l'ordine delle formule nell'antecedente o nel conseguente, o che permettono di evitare la ripetizione di formule ecc. Gentzen distingue infatti per i calcoli **L** gli schemi di *inferenza operativa* che regolano appunto la manipolazione degli operatori, dagli schemi di *inferenza strutturale* che invece consentono interventi del secondo tipo sulle sequenze. Diamo qualche esempio per i primi (le lettere maiuscole greche indicano succes-

sioni finite di formule, separate da virgole, eventualmente vuote):

1) *Introduzione di* $\vee$
nell'antecedente — nel conseguente

$$\frac{\mathscr{A}, \Gamma \vdash \Theta, \ \Gamma \vdash \Theta}{\mathscr{A} \vee \mathscr{B}, \Gamma \vdash \Theta} \qquad \frac{\Gamma \vdash \Theta, \mathscr{A}}{\Gamma \vdash \Theta, \mathscr{A} \vee \mathscr{B}} \qquad \frac{\Gamma \vdash \Theta, \mathscr{A}}{\Gamma \vdash \Theta, \mathscr{B} \vee \mathscr{A}}$$

2) *Introduzione di* $\neg$
nell'antecedente — nel conseguente

$$\frac{\Gamma \vdash \Theta, \mathscr{A}}{\neg \mathscr{A}, \Gamma \vdash \Theta} \qquad \frac{\mathscr{A}, \Gamma \vdash \Theta}{\Gamma \vdash \Theta, \neg \mathscr{A}} .$$

Per quanto riguarda gli schermi strutturali, oltre a quelli sopra ricordati informalmente, vogliamo esplicitarne uno che è decisamente centrale in tutto lo sviluppo del discorso di Gentzen. Si tratta dello schema di *taglio* (tedesco, *Schnitt*; inglese, *cut*), che si può porre sotto la forma:

$$\frac{\Gamma \vdash \mathscr{A}, \Theta \qquad \mathscr{A}, \Delta \vdash \Lambda}{\Gamma, \Delta \vdash \Theta, \Lambda}$$

e che si può leggere come segue: se in una derivazione figura una sequenza come quella in alto a sinistra (1) e quindi una come quella in alto a destra (2) che hanno in comune la *formula* $\mathscr{A}$, ma una nel conseguente e l'altra nell'antecedente, allora si può passare a una sequenza che ha come antecedente tutte le formule comprese nell'antecedente di (1) e di (2) meno la formula $\mathscr{A}$ e analogamente per il conseguente: in altri termini si può eliminare, tagliare, (lo *Schnitt* originale) la formula $\mathscr{A}$ in questione. Si noti che fra tutti gli schemi strutturali questo è l'unico che permette di eliminare completamente da una derivazione una formula (nel nostro caso $\mathscr{A}$; la formula così eliminata viene detta *formula di taglio*).

Ora i risultati cui prima si accennava relativi ai calcoli **L** riguardano appunto il ruolo di questo schema. Il primo di essi, noto come *Hauptsatz* (teorema principale) di Gentzen, afferma che ogni derivazione nei calcoli **L** di una sequenza $\Gamma \vdash \Theta$ può essere tra-

sformata in una derivazione della stessa sequenza nella quale *non figurino tagli*; in altri termini l'*Hauptsatz* afferma che tale schema è *superfluo*.[23] Allora dalla proprietà prima ricordata del taglio discende subito che dal momento che da una derivazione senza *tagli* non si può più eliminare alcuna formula, le formule di tutte le sequenze che intervengono nella derivazione di una sequenza data dovranno in qualche modo figurare in quest'ultima. È questo infatti un immediato corollario dell'*Hauptsatz*, noto come *proprietà o teorema della sottoformula* (*Teilformelsatz*): ogni derivazione nei calcoli **L** di una sequenza $\Gamma \vdash \Theta$ può essere trasformata in una analoga derivazione della stessa sequenza, tale che ogni formula occorrente in qualche sequenza della derivazione è sottoformula di almeno una formula di $\Gamma \vdash \Theta$. Esprimendoci intuitivamente, è sempre possibile derivare una sequenza senza ricorrere a formule «estranee» alla sequenza stessa: in un certo senso quindi la derivazione ha «memoria». Conviene ricordare qualche altra conseguenza immediata dello *Hauptsatz*: per prima cosa ne deriva la *coerenza* della logica predicativa classica e intuizionista. Allo scopo infatti basta far vedere che non è derivabile la sequenza vuota «$\vdash$»; ma ciò è immediato dal momento che l'unico modo di «far sparire» formule da una derivazione è quello di applicare il taglio, il che è appunto evitabile grazie al teorema principale di Gentzen. Ulteriori conseguenze estremamente significative per la logica intuizionista: una procedura di decisione per la logica proposizionale e una nuova dimostrazione della non derivabilità della legge del terzo escluso.[24]

[23] Data l'evidente analogia della regola del taglio con il *modus ponens*, questo risultato ne richiama uno del tutto analogo ottenuto, come si ricorderà, da Herbrand. Gentzen in effetti vede il collegamento in questione, ma per una errata valutazione del risultato di Herbrand ritiene di aver dimostrato un teorema assai più generale. In effetti, per quanto riguarda la logica classica i due teoremi sono sostanzialmente equivalenti. Il vantaggio tuttavia di quello ottenuto nella forma di Gentzen è che esso è applicabile (oltre che, ovviamente, come abbiamo visto, alla logica intuizionista) anche a molti altri sistemi non classici, ai quali invece non può essere esteso il teorema di Herbrand. Si veda comunque anche la successiva nota.

[24] Per la sola logica classica, ossia *per il solo* **LK**, l'*Hauptsatz* può essere rafforzato nel cosiddetto *verschärfter Hauptsatz*, secondo il quale la derivazione di una sequenza $\Gamma \vdash \Theta$ le cui formule siano tutte in forma normale prenessa può venir trasformata in una derivazione della stessa sequenza, che gode delle seguenti proprietà: 1) non contiene tagli; 2) contiene una sequenza «mediana» tale che 2.1)

L'introduzione dei calcoli **L** e in particolare la dimostrazione dell'*Hauptsatz* permettono a Gentzen di affrontare e risolvere il problema centrale di tutta la *Beweistheorie*: la dimostrazione della coerenza dell'aritmetica, che egli propone solo un anno dopo le *Ricerche*, in *Die Widerspruchfreiheit der reinen Zahlentheorie* (*La noncontraddittorietà della teoria dei numeri pura*, 1936), e ripropone modificata dopo qualche anno in *Neue Fassung des Widerspruchfreiheitsbeweises für die reine Zahlentheorie* (*Nuova versione della dimostrazione di noncontradditorietà per la teoria dei numeri pura*, 1938); importante anche, in questa connessione, l'articolo *Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie* (*Dimostrabilità e non dimostrabilità di forme ristrette dell'induzione transfinita nella teoria dei numeri pura*, 1943).

Non è qui possibile entrare nei particolari della dimostrazione di Gentzen. Ci limiteremo ad accennare molto in generale all'idea centrale: Gentzen associa ad ogni dimostrazione formale un numero ordinale transfinito $\alpha < \varepsilon_0$, che in un certo senso misura il «numero» delle applicazioni della regola del taglio, dove $\varepsilon_0 = \omega + \omega^\omega + \omega^{\omega^\omega} + \ldots$ Quindi fa vedere che, ove esistesse una dimostrazione di una formula contraddittoria, ad esempio della formula $1 \neq 1$, di ordinale, poniamo, β eliminando successivamente le applicazioni della regola del taglio si avrebbe anche una dimostrazione di $1 \neq 1$ cui sarebbe associato un numero ordinale $\gamma < \beta$ (intuitivamente: si avrebbe una dimostrazione «meno complessa» di $1 \neq 1$); ne segue che l'esistenza di una contraddizione contrav-

né tale sequenza né la sua derivazione contengono quantificatori e 2.2) nel resto della derivazione vengono impiegati soltanto schemi di inferenza strutturale e schemi relativi ai quantificatori. In altri termini, la sequenza mediana divide per così dire la derivazione in due parti, una puramente logico-proposizionale, l'altra puramente logico-predicativa (si noti, in collegamento al teorema di Herbrand, che la sequenza mediana corrisponde appunto alla formula di Herbrand). Questa forma rafforzata dell'*Hauptsatz* viene sfruttata da Gentzen, sempre nelle *Ricerche*, per dare una nuova dimostrazione della coerenza dell'aritmetica *senza* induzione (come si ricorderà per l'aritmetica senza induzione o con l'induzione limitata a formule di tipo particolare altre dimostrazioni erano state date da Ackermann e von Neumann nel 1927, da Herbrand nel 1932; e un teorema generale al riguardo viene dato da Bernays nel 1936). Tuttavia, come lo stesso Gentzen concludeva, «l'aritmetica senza induzione completa è... di scarso significato pratico, perché l'induzione completa è costantemente richiesta nella teoria dei numeri. Finora peraltro non è stata dimostrata la coerenza dell'aritmetica con induzione completa». È appunto quest'ultimo l'importante ulteriore risultato raggiunto da Gentzen nell'anno successivo.

verrebbe al principio di induzione transfinita (fino a ε_0) o, in termini meno esatti, si potrebbe già svolgere nel «finito».

È ovvio che il metodo al quale si appella Gentzen per dare la sua dimostrazione di coerenza trascende in qualche senso (come minimo appunto per il teorema di Gödel) quanto è «formalizzabile» nell'aritmetica, ossia va al di là dell'ambito «finitista» in senso hilbertiano.[25]

Ma, afferma Gentzen, «se si deve eseguire una prova di coerenza, ovviamente nella dimostrazione devono essere usate certe forme di inferenza. La correttezza di *queste* inferenze deve essere pre-

[25] Si consideri il principio di induzione matematica posto nella forma (cui in particolare si dà il nome di induzione completa)

$$\forall y \{\forall x(x < y \rightarrow P(x)) \rightarrow P(y)\} \rightarrow \forall x P(x) \tag{1}$$

ossia: *se* per ogni numero naturale y, posto che tutti i numeri naturali minori di y godano di una certa proprietà P, allora anche y gode di P, *allora* tutti i numeri naturali godono della proprietà P. Se in questa formulazione si pone ovunque «numero ordinale» al posto di «numero naturale» e «che precedono y in un buon ordine $<$» al posto di «minori» si ha un'enunciazione del principio di induzione transfinita. Si dimostra facilmente che quest'ultimo è applicabile a tutti e soli gli insiemi bene ordinati. Estendendosi su tutti gli ordinali (e non solo sugli ordinali finiti, come l'induzione completa) questo principio – anche dal punto di vista intuitivo – è estremamente «forte»: una sua adeguata formulazione e giustificazione è infatti possibile solo nella teoria degli insiemi. Ora Gentzen trova una opportuna formula $x < y$ *esprimibile* nell'aritmetica, che definisce un buon ordine dei naturali di tipo ε_0. Fa inoltre vedere che l'induzione transfinita limitata per così dire al buon ordine $<$, ossia «troncata» a ε_0 (si osservi che nel caso dell'induzione completa ci si spinge fino a ω) permette, nel senso accennato nel testo, di dimostrare la coerenza dell'aritmetica. Anzi Gentzen fa anche vedere – tenuto conto del teorema di Gödel – che questa scelta del buon ordine $<$ è ottimale, nel senso che se ci si limita a considerare un *segmento* $<_1$ di $<$ (ossia se si spinge l'induzione fino a un ordinale minore di ε_0) il principio di induzione transfinita per $<_1$ è *dimostrabile* nell'aritmetica. È quasi superfluo aggiungere che, per il teorema di Gödel, ciò non può avvenire per l'induzione fino a ε_0 (in termini diversi, non è *dimostrabile* in $\mathfrak{P}$ la formula che esprime l'induzione su $<$ di tipo ε_0). È forse più chiaro ora capire il ruolo della induzione fino a ε_0 nella dimostrazione del teorema. Supponiamo infatti di avere ottenuto una derivazione di una contraddizione; riduciamo tale derivazione in modo effettivo, costruttivo, a una derivazione che abbia la stessa conclusione ma numero ordinale minore eliminando i tagli. Questa procedura termina dopo un numero finito di passi e quindi si ottiene una dimostrazione non più riducibile, cioè che non contiene più applicazioni della regola del taglio ed è facile verificare che dimostrazioni di questo tipo non possono terminare con $1 \neq 1$. Come si nota la tecnica è la stessa del caso dell'*Hauptsatz* per i calcoli logici solo che qui – in vista della presenza del principio d'induzione – occorre utilizzare non più l'induzione fino a ω ma quella fino a ε_0.

supposta dall'inizio, altrimenti tutta la dimostrazione risulterebbe ovviamente circolare. Non può esistere una dimostrazione "assoluta" di coerenza... Le inferenze devono essere compatibili con il punto di vista *costruttivo*, l'affidabilità del quale non è posta in discussione ma data per garantita. L'obiettivo è allora quello di dimostrare la coerenza dell'interpretazione attualista per mezzo di inferenze costruttive».

L'accettare o meno certe forme di inferenza come «finitiste» non è legato alla vaga intuizione di «finito» come sembrava intendere, pur senza averlo mai precisato, Hilbert; ma a una convenzione iniziale, anch'essa del resto di tipo non formale, che si adegui a una concezione costruttiva, potenziale dell'infinito. Naturalmente non si può *dimostrare* una tale adeguazione «se non altro perché la nozione di "finitista" non è univocamente definita da un punto di vista formale e in effetti non può essere delimitata in questo modo. Tutto quello che possiamo fare è esaminare ogni singola inferenza da questo angolo visuale e tentare di accertare se quell'inferenza è in armonia col senso finitista dei concetti che vi intervengono e assicurarsi che essa non si fondi su un'inammissibile interpretazione "attualista" di questi concetti».

Per dare un senso concreto a questo discorso e togliere l'impressione che Gentzen pretenda di giustificare l'induzione fino ad ω ricorrendo all'induzione su ε_0 (un *ordinale* ben più grande di ω!) è essenziale osservare che la dimostrazione di Gentzen coinvolge solo proprietà finitiste, combinatorie, delle dimostrazioni e solo per queste proprietà (non per proprietà arbitrarie come nello schema di induzione su ω postulato nell'aritmetica) viene utilizzata l'induzione su ε_0. In altre parole, Gentzen ragiona su oggetti concreti e proprietà concrete facendo riferimento ad un buon ordine la cui fondatezza non è giustificabile nell'aritmetica. È solo questo l'elemento «infinitario» nella dimostrazione di Gentzen che per il resto procede in modo puramente combinatorio.

È proprio qui, in definitiva, che Gentzen prende le distanze da Hilbert da una parte e dagli intuizionisti dall'altra. Dal primo, precisando la nozione di «finitista», collegandola sostanzialmente al segmento di ordinali lungo il quale si intende sviluppare l'induzione cosicché, riprendendo il punto da cui eravamo partiti, «l'aumento di complessità del concetto di infinito ai tre livelli della matematica... teoria elementare dei numeri, Analisi e teoria degli in-

siemi, è accompagnato da una corrispondente estensione della sequenza dei numeri ordinali transfiniti...». Il significato del risultato di Gentzen non fu immediatamente compreso, sicché non fu subito visto come un contributo al programma hilbertiano come concepito originariamente; tuttavia, come afferma Solomon Feferman, necessariamente «le limitazioni teoriche imposte a quel programma dal teorema di Gödel avrebbero condotto a considerare una *estensione del programma di Hilbert* nella quale la richiesta che tutti i ragionamenti della teoria della dimostrazione fossero finitisti sarebbe stata indebolita richiedendo che tutti questi ragionamenti fossero costruttivi». E ciò spiega il grande impatto che i metodi e i risultati di Gentzen hanno avuto e continuano ad avere nelle ricerche attuali di teoria della dimostrazione.[26]

Dagli intuizionisti Gentzen invece si discosta perché, lungi dal rifiutare come prive di senso le proposizioni matematiche transfinite (ideali in senso hilbertiano) ritiene necessario tentare di dar loro un «senso finitista» (anche se nella sua accezione allargata) evitando così mutilazioni catastrofiche della matematica: «La maggior parte della mia dimostrazione di coerenza», afferma Gentzen, «... consiste proprio nell'*ascrivere un senso finitista* a proposizioni attualiste, e precisamente: per ogni proposizione arbitraria, nella misura in cui è dimostrabile, può *essere enunciata*... una *re-*

[26] È opportuno anticipare già a questo punto altre proposte successive al lavoro di Gentzen, avanzate per risolvere il problema della coerenza dell'aritmetica e di sistemi più forti (Analisi), di cui avremo occasione di parlare più avanti. Va detto intanto che Hilbert-Bernays nel 1939 non solo sembrano ammettere l'estensione di Gentzen dei loro metodi, ma anzi sembrano ritenere che estendendo ulteriormente tali metodi si possa giungere a una dimostrazione della coerenza dell'Analisi. In effetti tali estensioni si sono avute e sono state finora sostanzialmente di due tipi. Da una parte Gaisi Takeuti nel 1953 ha proposto un calcolo logico generale (del tipo dei calcoli di Gentzen ma per la teoria dei tipi con variabili predicative di ordine superiore) e ha congetturato che per tali calcoli valga ancora un analogo dell'*Hauptsatz*. Se tale congettura si fosse rivelata esatta e fosse stata suscettibile di dimostrazione *costruttiva*, ne sarebbe seguita una dimostrazione di coerenza dell'Analisi. Vedremo più avanti come la congettura sia stata affrontata dalla ricerca successiva. L'altro aspetto delle estensioni riguarda in modo naturale la definizione costruttiva di ordinali maggiori di ε_0 e tali che con l'induzione spinta fino ad essi si possa dimostrare la coerenza di varie teorie matematiche; ricordiamo in questo campo i lavori e i risultati di Ackermann (1951), Kurt Schütte (1960), ancora Takeuti (1956) che riprendono precedenti ricerche di O. Veblen e S.K. Kleene e A. Church sulle notazioni ordinali. Delle proposte di Gödel del 1958 parleremo più avanti.

gola di riduzione, e ciò rappresenta il senso finitista della proposizione stessa e questo senso si acquisisce proprio per mezzo della prova di coerenza.»

A mo' di conclusione, riportiamo le parole con cui Andrzej Mostowski valuta congiuntamente i risultati di Herbrand e Gentzen, accomunandoli nella concezione finitista: «Esistono indubbiamente due opposte tendenze nello studio dei fondamenti della matematica: l'infinitistica o insiemistica e la finitistica o aritmetica. Le scoperte originali di Herbrand e di Gentzen appartengono ovviamente alla seconda di queste tendenze, ma la ricerca successiva che si è fondata su questi risultati ha tratto molte idee dalla prima. L'influenza dell'approccio insiemistico è chiaramente riscontrabile nel teorema di coerenza di Bernays[27] nel quale le nozioni semantiche sono volutamente imitate in termini finitisti. Possiamo dire che i metodi di Herbrand e di Gentzen ci permettono di rendere finitisti certi casi particolari di costruzioni insiemistiche.» È proprio in questo loro ruolo che le tecniche della *Beweistheorie* si pongono come tentativo sistematico di enucleare il significato algoritmico finito dalla matematica astratta.

6. LA PRECISAZIONE DEL CONCETTO DI «EFFETTIVO»: LA TEORIA DELLA RICORSIVITÀ

Come ricordato a suo tempo, la chiarificazione matematica delle nozioni di calcolabilità effettiva e di decidibilità, costituisce una delle maggiori conquiste della ricerca logica degli anni trenta, tanto sul piano matematico quanto su quello più genericamente logico e filosofico. I concetti di funzione computabile e di problema decidibile, strettamente legati l'uno all'altro, fanno da sfondo esplicito o implicito tanto all'idea skolemiana di una teoria dei nu-

[27] Si tratta del teorema dato da Bernays nel 1936 (si veda la nota 24 a pag. 619) e poi riportata nel secondo volume (1939) dei *Fondamenti della matematica* di Hilbert-Bernays. In esso Bernays dimostra che data una teoria i cui assiomi posti in forma prenessa abbiano prefisso del tipo ∀∃ (ossia ogni quantificatore universale precede ogni quantificatore esistenziale) se gli assiomi sono veri in modo «effettivo», allora lo sono anche i teoremi (da questo risultato segue immediatamente la coerenza dell'aritmetica con induzione ristretta a formule prive di quantificatori).

meri non fondata su dubbi concetti insiemistici, quanto al finitismo hilbertiano. Anche le concezioni costruttiviste privilegiano le funzioni e i concetti formulati in termini di costruzioni e algoritmi, ma è soprattutto con Hilbert che la calcolabilità e la decidibilità acquistano un ruolo centrale come paradigmi di sicurezza, aggancio al concreto e – in ultima istanza – criterio di significatività. Abbiamo già ricordato come Hilbert concepisse l'indagine metamatematica in termini di ricerca sulle condizioni di dimostrabilità e quindi di risolubilità di problemi. Questo atteggiamento – che trova la sua canonizzazione nella famosa conferenza tenuta da Hilbert a Parigi nel 1900, ove formulava alcuni dei problemi la cui risoluzione avrebbe costituito un passo avanti decisivo per la matematica – è presente già nei *Fondamenti della geometria* e rimanda come paradigma alla teoria di Galois sulla risolubilità per radicali delle equazioni polinomiali.

Come si ricorderà, di fronte a ogni problema, secondo Hilbert, il matematico deve domandarsi: È risolubile con i mezzi a disposizione? Se sì, come risolverlo e se no, quali altri mezzi permettono di farlo? In questa concezione del modo in cui la matematica progredisce, si intrecciano momenti in cui si procede in modo tradizionale sfruttando le ipotesi – i mezzi – di cui si dispone già, e momenti in cui deve intervenire l'analisi metamatematica per indagare se il problema posto è risolubile (ad esempio, se un dato enunciato è dimostrabile a partire dalle ipotesi aggiunte) oppure se è compatibile assumere altre ipotesi e con esse risolvere il problema. La questione dei fondamenti della matematica – vista come dimostrazione della coerenza delle teorie fondamentali e come prova che le ipotesi non concrete che hanno solo validità formale sono un'estensione conservativa rispetto agli enunciati con contenuto materiale – è quindi solo un aspetto di una problematica più generale riguardante le teorie formalizzate, le uniche su cui può agire concretamente l'indagine metamatematica. Il problema dei problemi diviene così per Hilbert la questione di decidere per ogni teoria $\mathfrak{T}$ quali sono i suoi teoremi. Come sappiamo è questo il cosiddetto *Entscheidungsproblem* che costituisce il nucleo della *Beweistheorie* hilbertiana e della metamatematica finitista nel suo complesso.

La problematica non è solo logica ma riflette una dimensione della ricerca matematica che sotto lo stimolo di Hilbert si va facendo sempre più precisa negli anni che vanno dall'inizio del seco-

lo fino alla fine del 1940. Si tratta di indagare il significato computazionale, la dimensione algoritmica di nozioni che la matematica astratta (quel «paradiso di Cantor» che proprio Hilbert era deciso a difendere con i denti) andava via via sviluppando. Proprio nel 1911 Max Dehn (1878-1952), che aveva risolto il quarto dei problemi presentati da Hilbert, getta le basi della moderna teoria combinatoria dei gruppi vista come analisi delle interrelazioni tra presentazioni concrete, linguistiche, dei gruppi e gruppi astratti, proponendo all'attenzione dei matematici i tre problemi fondamentali: quello della parola (quando due termini di una presentazione, due parole, determinano lo stesso oggetto nel gruppo presentato?) quello della coniugazione (quando due parole denotano elementi coniugati?) e infine quello dell'isomorfismo (quando due presentazioni danno origine a gruppi isomorfi?). Dehn mostrava non solo come questi problemi riflettessero questioni basilari sui gruppi fondamentali e quindi sulle proprietà topologiche degli spazi associati, ma anche come essi riguardassero l'esistenza o meno di algoritmi per decidere o per calcolare.

Non c'era in sostanza grande differenza tra un problema come quello della parola (che si può tradurre in quello di sapere quando da date relazioni definitorie, le equazioni di una presentazione, si può dedurre una data equazione tra due parole) ed il problema all'apparenza squisitamente logico di decidere quali siano i teoremi di una teoria formalizzata. Allo stesso modo – come avrebbero verificato di lì a poco matematici come O. Schreier (1901-29) e E. Artin (1892-1962) – c'era pure una forte analogia tra i problemi di forma normale in logica e nella teoria combinatoria dei gruppi, come anche tra il problema della noncontraddittorietà delle teorie e quello della esistenza di gruppi che soddisfano date presentazioni, come succede ad esempio definendo il prodotto amalgamato.

Discorsi analoghi negli anni trenta avrebbe sviluppato Grete Hermann, allieva di Noether, nello studio delle basi degli ideali e della teoria della divisibilità, riprendendo il sogno kroneckeriano di una fusione di teoria dei numeri e geometria su base combinatoria, in contrapposizione a quella dedekindiana di tipo astratto fondata sulla teoria degli ideali. L'interesse per problemi di tipo combinatorio e la conseguente idea di analizzare in astratto l'idea di computo mediante regole è quindi simultanea

allo svilupparsi della matematica astratta ed ha tanto motivazioni strettamente tecniche (algebra costruttiva, teoria combinatoria dei gruppi) quanto filosofiche. Queste ultime non riguardano soltanto costruttivisti e finitisti come Borel (che già nel 1913 tentava una definizione generale di algoritmo) o Hilbert o Skolem, ma anche pensatori di ispirazione molto diversa come E. Post, che si muoveva su uno sfondo bergsoniano ed era interessato a indagare la zona d'ombra tra procedura meccanica e libera attività psichica.

Il nesso tra tutti questi problemi è costituito dal concetto di algoritmo, anche se non tutte le precisazioni del concetto di effettivo date negli anni trenta assumono questo concetto come primitivo. Al termine di algoritmo di decisione o di computo – lasciamo per ora indeterminato – si associa di solito intuitivamente l'idea di un complesso di istruzioni tali che: *a*) siano precisamente determinate sì da non consentire situazioni di dubbio nel corso dell'esecuzione, e inoltre «universalmente» comprensibili nel senso che chiunque possa applicarle una volta presane conoscenza; in altri termini un algoritmo deve essere «deterministico» e essenzialmente «meccanico»; *b*) siano abbastanza generali da potersi applicare a ogni problema di una data classe; *c*) applicate a certi «dati» forniscano criteri per determinare quando la soluzione è raggiunta, e questo avvenga dopo un numero finito di passi. Nel caso la soluzione non venga raggiunta, le istruzioni possono dar luogo a un processo che prosegue indefinitamente.

Esempi a tutti noti di procedure di questo tipo sono l'algoritmo euclideo per la ricerca del massimo comune divisore fra due numeri naturali; o il semplicissimo algoritmo col quale stabilire se un qualunque numero naturale dato è primo o no; l'algoritmo (tavole di verità) mediante il quale stabilire se una data formula proposizionale è o no una tautologia (che risolve il *problema della decisione* per la logica proposizionale); ecc. Abbiamo tuttavia ricordato altri problemi per i quali non conosciamo un algoritmo in grado di offrirci una soluzione, ad esempio il problema della decisione per la logica del primo ordine **LP**: data una qualunque formula $\mathscr{A}$ scritta nel linguaggio $\mathscr{L}$ di **LP**, è possibile decidere in un numero finito di passi se $\mathscr{A}$ è o no un teorema di **LP**? Si noti che questa domanda equivale alla seguente: considerata la classe C di tutte le formule **LP**, è possibile dare un algoritmo per isola-

re in essa una sottoclasse T costituita da tutte e sole quelle formule che sono teoremi di **LP**? Ovvero è possibile *decidere* per ogni formula $\mathscr{A}$ di **LP** se $\mathscr{A} \in T$ oppure no? Il lettore potrebbe pensare che in effetti tale algoritmo esista e sia proprio rappresentato da un sistema di assiomi per **LP** con le regole di derivazione a suo tempo viste; infatti è chiaramente possibile «decidere» per ogni formula $\mathscr{A}$ di **LP** se essa è o no un assioma e d'altra parte, data una successione finita di formule, siamo in grado di «decidere» se essa è una derivazione in **LP** (in altri termini, proprio per la definizione formalista di sistema formale anche le regole di derivazione sono «decidibili»).

Ma questo non basta: per determinare quando una data formula è un teorema dobbiamo essere in grado di sapere se *esiste* una sua dimostrazione ed il dominio delle dimostrazioni è un insieme infinito (anche se decidibile); il carattere «effettivo» delle regole e della proprietà di essere assioma ci permette sì di determinare, *data* una successione di formule, se essa è una dimostrazione, ma non ci consente in alcun modo, data una qualunque formula $\mathscr{A}$, di sapere *se esiste* una successione di formule che sia una dimostrazione di $\mathscr{A}$. In altri termini, data una formula, non è detto che si abbia in generale un modo di ricostruire le sue possibili derivazioni. La situazione può essere cioè descritta intuitivamente dicendo che l'insieme dei teoremi di **LP** è *effettivamente generabile* ma non è detto a priori che sia *decidibile*: questa distinzione introduce due possibili direzioni di approccio, storicamente seguite, come vedremo, per una precisazione rigorosa del problema dell'effettivo. Si vedrà infatti che i due concetti sono tra loro interdefinibili ma non riducibili l'uno all'altro. È chiaro comunque che la precisazione rigorosa di entrambi questi concetti dipende da una previa precisazione del concetto di algoritmo, al quale tutti e due rimandano con la clausola di «effettività».

Orbene, finché ci limitiamo alla caratterizzazione intuitiva di algoritmo sopra data, potremo fare solo enunciazioni di tipo per così dire positivo: «per tale problema (di una certa classe di problemi) esiste un algoritmo fatto così e così» ed è in questo senso – come si è detto – che erano stati affrontati i problemi di decisione fino agli anni trenta; ma è impossibile ottenere risultati *negativi*, ossia far vedere che un determinato algoritmo richiesto *non* esiste se ci si limita ad una descrizione così vaga del concetto stesso. A

partire allora da questa accezione intuitiva si impone una sua precisazione rigorosa; è appunto a questa precisazione che si rivolge negli anni trenta la teoria della ricorsività. Prima di presentare, molto brevemente alcune delle soluzioni proposte in quegli anni, consideriamo alcuni concetti intuitivi, e i loro rapporti, che ci serviranno nel nostro discorso. Ricordando l'esempio sopra fatto a proposito del problema della decisione per **LP**, diciamo allora *decidibile* un insieme M se è possibile dare un algoritmo che permetta di decidere per ogni oggetto x se $x \in M$ oppure $x \notin M$. Consideriamo ora un predicato a n posti; è chiaro che potremmo riguardarlo da un punto di vista *estensionale* come l'insieme di tutte e sole quelle n-uple di oggetti che godono della proprietà espressa del predicato (useremo convenzionalmente *attributo* per indicare questa accezione estensionale di un predicato). Allora viene naturale chiamare *decidibile* un certo predicato a n posti se è decidibile nel senso detto sopra l'attributo corrispondente, ossia se per ogni n-upla di oggetti esiste un algoritmo che permette di stabilire se tale n-upla appartiene o no all'attributo. Passiamo ora agli algoritmi di computazione e consideriamo una funzione a n argomenti: diremo effettivamente *computabile* (o calcolabile) una tale funzione se, data comunque una n-upla di suoi argomenti, è possibile (esiste un algoritmo per) trovare in un numero finito di passi il valore corrispondente. Decidibilità e computabilità sono collegate: sappiamo infatti che a ogni insieme M (e quindi in particolare a ogni attributo) è associabile una funzione, precisamente la sua *funzione caratteristica* f_M e dovrebbe allora risultare chiaro che un insieme M è *decidibile* se e solo se la sua funzione caratteristica f_M è *computabile.*

Risulta così (e il lettore può facilmente convincersene) che *essere computabile* di una funzione, *essere decidibile* di un insieme o di un predicato, e *possedere un algoritmo* sono concetti fra loro interdefinibili; e che quindi per dare una precisazione rigorosa del concetto intuitivo di *procedura effettiva*, basta riferirsi a uno solo di essi indifferentemente. Ci si convince anche facilmente che se l'insieme M è decidibile, tale è anche il suo complemento $\overline{M}$, *e che se* M, N sono decidibili, tali saranno anche gli insiemi $M \cap N$, $M \cup N$. Ancora, se il predicato $P(x_1, ..., x_n)$ è decidibile, tale sarà anche il predicato $\neg P(x_1, ..., x_n)$, sua negazione; e che se $P(x_1, ..., x_n)$ e $Q(x_1, ..., x_n)$ sono decidibili, tali saranno anche i predicati $P(x_1, ..., x_n)$ **o** $Q(x_1,$

..., x_n), dove **o** venga sostituito da uno qualunque dei connettivi, $\wedge, \vee, \rightarrow, \leftrightarrow$; in particolare come già mostrato da Skolem sarà decidibile il predicato

$$P(x_1, \ldots, x_n, y) = \overset{y}{\underset{x=0}{\exists}} x Q(x_1, \ldots, x_n, x),$$

ottenuto per quantificazione esistenziale *limitata* dal predicato decidibile $Q(x_1, \ldots, x_n, x)$ e analogamente per la quantificazione universale *limitata*. Per l'altro concetto intuitivo che abbiamo introdotto, quello di insieme *effettivamente generabile*, da quanto sopra detto risulta chiaramente che tale è un insieme M per il quale esiste un procedimento effettivo, un algoritmo, che *genera* uno dopo l'altro tutti gli x tali che $x \in M$; (in altri termini si considera effettivamente generabile un insieme che sia campo di valori di una funzione computabile). Risulta allora chiaro che il concetto di insieme effettivamente generabile non coincide *a priori* con quello di insieme decidibile: in questo secondo caso, dato un oggetto qualunque, sappiamo decidere se esso appartiene o no all'insieme; nel primo caso, invece, tutto quello che sappiamo è che *se* un certo x appartiene all'insieme esso verrà «generato» a un certo punto del procedimento effettivo. Tra i due concetti esiste tuttavia una certa connessione, che è la seguente: un insieme M effettivamente generabile è decidibile se e solo se anche il suo complemento $\overline{M}$ è effettivamente generabile. È questo il contenuto di un teorema di E. Post, ed è facile rendersene conto: se è possibile generare effettivamente tanto gli $x \in M$ quanto gli $x \in \overline{M}$ (ossia gli x che $\notin M$) è chiaro che procedendo a zig zag in un numero finito di passi potremo decidere a quale dei due insiemi M o $\overline{M}$ un dato oggetto appartiene.

Ancora un'osservazione preliminare. Sinora abbiamo parlato di algoritmi senza specificare i domini sui quali essi sono definiti, senza cioè dire che *tipo di oggetti* questi algoritmi manipolano. È chiaro che si dovrà trattare di oggetti concreti e che anche se si vuole disporre di algoritmi che riguardano oggetti astratti (ad esempio, numeri naturali), la definizione di algoritmo dovrà essere data modulo una previa determinazione *linguistica*, di nomi (oggetti concreti, appunto) assegnati biunivocamente agli enti astratti. In altri termini un algoritmo dovrà es-

sere sempre definito relativamente ad un linguaggio finitamente specificato (ossia con alfabeto finito e con regole finite per la costruzione delle espressioni) e quindi con un insieme di espressioni al più numerabile. D'altra parte, mediante la tecnica dell'aritmetizzazione, sarà possibile in certo senso *normalizzare* ogni linguaggio di questo tipo assegnando biunivocamente e in modo effettivo a ogni espressione del linguaggio un *numero naturale* (il suo numero di Gödel): questo intanto spiega perché in generale ci si possa limitare a considerare funzioni numeriche (ossia che prendono argomenti e valori su N) predicati numerici (ossia che hanno come argomenti n-uple di numeri naturali e come valori uno dei due valori di verità Vero o Falso) insiemi (e in particolare attributi) di numeri naturali ecc. Ogni algoritmo si potrà allora considerare come descrivente una funzione su *numeri*; ma proprio per quanto detto sopra, dovremo servirci, per la definizione concreta dell'algoritmo stesso, di un linguaggio che possegga *nomi* per i numeri, ossia *cifre* le quali, si noti bene, non dovranno necessariamente essere quelle ordinarie della notazione decimale, ma dipenderanno da considerazioni generali legate al tipo di approccio che si dà al problema della caratterizzazione del concetto di algoritmo. Il lettore non si dovrà quindi stupire se nel seguito, parlando delle definizioni proposte da Turing, Church, Post, Markov parleremo di «funzioni numeriche» e di relativi algoritmi anche se, almeno in modo immediato, queste definizioni non sembrano riguardare i numeri: i linguaggi in questione, malgrado le apparenze, contengono tutti espressioni particolari che fungono da *cifre* (ossia, ripetiamolo, nomi per i numeri naturali).[28]

Il primo passo consapevole verso una precisazione rigorosa di tutto questo bagaglio intuitivo – dopo le anticipazioni di Skolem cui abbiamo già accennato ed analoghe indicazioni di Herbrand – venne compiuto come sappiamo da Gödel nella memoria del 1931. Gödel individuava una classe di funzioni numeriche, oggi dette *ricorsive primitive*, sul cui carattere di effettiva computabilità

[28] Come già in precedenza, noi indicheremo genericamente con $\bar{n}$ la *cifra* del *numero n*. Si osservi ancora che quando si vuol parlare della *cifra* corrispondente a un certo *numero di Gödel*, si usa spesso il termine *gödeliano*.

non potevano sorgere dubbi.[29] Infatti tali funzioni vengono definite assumendo alcune funzioni di base certamente computabili da un punto di vista intuitivo, e quindi esplicitando due schemi di formazione che a partire da funzioni ricorsive primitive permettano di definire altre funzioni ancora ricorsive primitive. Le funzioni base sono:

1) la *funzione di successore* $S(x) = x + 1$
che applicata a un qualunque numero naturale dà come risultato il successore di quel numero (tale funzione è ovviamente computabile);

2) la *funzione di costante* 0, $C(x) = 0$
ossia quella funzione, anch'essa banalmente computabile, che ad ogni argomento fa corrispondere come valore lo 0;

3) le *funzioni di selezione* $U_i^n(x_1, ..., x_n) = x_i$ per ogni n finito e per ogni i $(1 \leq i \leq n)$; queste funzioni applicate a una qualunque n-upla di argomenti le fanno corrispondere come valore l'i-esimo elemento della n-upla; anch'esse sono banalmente computabili.

I due schemi di formazione sono i seguenti, dove $g, h_1, ..., h_m, h'$, g' sono funzioni ricorsive primitive

SS $\qquad f(x_1, ..., x_n) = g(h_1(x_1, ..., x_n), ..., h_m(x_1, ..., x_n));$

[29] Si osservi che Gödel non intendeva assolutamente identificare questa classe di funzioni con la classe delle funzioni effettivamente computabili; del resto già nel 1928 Ackermann in *Zum Hilbertschen Aufbau der reellen Zahlen* (*Sulla costruzione hilbertiana dei numeri reali*) aveva esibito una funzione che, pur essendo computabile, non risultava ricorsiva primitiva. Il senso della proposta di Gödel è un altro: le funzioni ricorsive primitive costituiscono una porzione (molto ristretta) di aritmetica sufficiente alla formalizzazione della sintassi di una teoria. Di qui il collegamento con i lavori di Skolem e la loro interpretazione da parte di Hilbert. Da quanto sappiamo circa la conduzione della dimostrazione di Gödel, possiamo dire che *tutte* le funzioni ricorsive primitive sono rappresentabili formalmente (nel senso della nota 7 a pag. 569) in $\mathfrak{P}$; vedremo tuttavia che esse non sono le *sole* rappresentabili in questo senso, ossia la classe delle funzioni formalmente rappresentabili in $\mathfrak{P}$ è più ampia, comprende in modo proprio la classe delle funzioni ricorsive primitive. Sarà proprio da un'analisi del concetto di rappresentabilità formale, come vedremo, che prende lo spunto una caratterizzazione (quella proposta da Herbrand, Gödel, Kleene) della classe delle funzioni effettivamente computabili.

SR $$\begin{cases} f(0, x_1, ..., x_n) = h'(x_1, ..., x_n) \\ f(x + 1, x_1, ..., x_n) = g'(x, f(x, x_1, ..., x_n), x_1, ..., x_n). \end{cases}$$

I due schemi precedenti, detti rispettivamente di *sostituzione* e di *ricorsione*, vanno letti come segue: *a*) si può definire una funzione f a n argomenti, a partire da una funzione data g, a m argomenti, sostituendo al posto degli argomenti stessi m funzioni $h_1, ..., h_m$ a n argomenti; e il suo senso è ovviamente che se, come supposto, g e h_1, ..., h_m sono *ricorsive primitive*, allora anche la funzione f è ricorsiva primitiva; *b*) si può definire una funzione f assegnando il suo valore per l'argomento 0 (e per eventuali altri *parametri* $x_1, ..., x_n$) e quindi dando il modo di calcolare il valore per ogni argomento successivo $x + 1$ in funzione dell'argomento precedente e del valore che la funzione stessa aveva assunto per x (oltre, ovviamente, in dipendenza dai soliti parametri). Si dice in tal caso che f è definita per *ricorsione* o per *induzione* (su x) donde il nome dello schema. Il lettore può facilmente verificare che gli assiomi 2.4 e 2.5 di $\mathfrak{P}$ (paragrafo III.1) altro non sono che le definizioni ricorsive delle funzioni di somma e prodotto; come pure tali sono le definizioni della nota 47 a pag. 529. Ovviamente se tutte le funzioni che intervengono in SR nelle equazioni a secondo membro sono ricorsive primitive – come supposto – allora anche la f così definita è ricorsiva primitiva. Agli schemi SS e SR sono riconducibili, come hanno mostrato successivamente Gödel, Rosza Peter (1905-77), Hilbert-Bernays e altri fra il 1931 e il 1934, ulteriori schemi di formazione per funzioni ricorsive primitive, ad esempio: lo schema di ricorsione sul decorso di valori, nel quale per la definizione di $f(x + 1)$ si può far ricorso a un numero arbitrario di valori assunti da f per argomenti minori di x (e non solo per l'immediato predecessore); la ricorsione multipla, nella quale si può operare contemporaneamente la ricorsione su più variabili; e ancora la distinzione dei casi, la minimalizzazione e altri sui quali è inutile soffermarsi in questa sede.

Che il concetto di funzione ricorsiva primitiva non sia un corrispettivo rigoroso adeguato a quello di funzione computabile è già stato detto; oltre all'esempio di Ackermann, nel 1935 la Peter, che fu uno dei pionieri nel campo e alla quale si deve – tra l'altro – il nome stesso di funzione ricorsiva, in *Konstruktion nicht rekursiver Funktionen* (*Costruzione di funzioni non ricorsive*) approfondiva notevolmente la questione. Ma pur senza presentare un esempio parti-

colare ci si può convincere facilmente in generale che esistono funzioni computabili e non primitive ricorsive. Una via è la seguente. Si dimostra facilmente che la classe delle funzioni ricorsive primitive è generabile in modo induttivo e che è possibile *enumerare effettivamente*, assegnando ad ogni funzione il numero di Gödel di una sua definizione in termini di funzioni iniziali e schemi, tutti i suoi elementi, ossia porli in una successione f_1, f_2, f_3, ... Definiamo ora una funzione $h(x)$ ponendo

$$h(x) = f_n(x) + 1$$

ossia per ogni n il valore di h per l'argomento x è uguale al valore in x dell'n-esima funzione della nostra enumerazione, aumentato di 1; poiché f_n è per ipotesi ricorsiva primitiva e quindi calcolabile, anche h ovviamente sarà calcolabile; se ora $h(x)$ fosse a sua volta ricorsiva primitiva, dovrebbe coincidere con una delle f_i; poniamo coincida con f_m. Si avrebbe allora, calcolando i valori proprio per l'argomento m

$$f_m(m) = h(m) = f_m(m) + 1$$

ossia una contraddizione.[30]

Come determinare allora *tutte* le funzioni computabili?

Fra il 1934 e il 1936 vi furono numerose proposte di caratterizzazione di una classe più generale di funzioni che potesse aspirare ad essere identificata con la classe delle funzioni computabili. Una prima proposta si ebbe nel 1934, da parte di Gödel, il quale in *On undecidable propositions of formal mathematical systems* (*Sulle proposizioni indecidibili dei sistemi matematici formali*), riprendendo un'idea avanzata da Herbrand nel 1931, proponeva di identificare le funzioni computabili con quelle che ora egli chiamava funzioni *ricorsive generali* e che noi diremo *ricorsive* senza ulteriori qualificazioni. Il procedimento di Gödel può così schematizzarsi. Fissato un sistema formale le cui formule sono costituite da equazioni tra termini

[30] Si noti l'analogia col processo di «diagonalizzazione» usato da Cantor per dimostrare la non numerabilità dell'insieme dei numeri reali. La funzione $h(x)$ così definita o, per essere più precisi, una funzione $F(n, x)$ che per ogni n assuma, per l'argomento x, il valore dell'n-esima funzione ricorsiva primitiva per lo stesso argomento, si dice *funzione universale* per la classe delle funzioni ricorsive primitive.

e le cui regole permettono di passare da un'equazione a un'altra mediante sostituzioni e rimpiazzamenti opportuni, e definiscono quindi un *calcolo di equazioni*, funzione ricorsiva è per definizione ogni funzione $f(x_1, ..., x_n)$ per cui esiste nel calcolo un sistema di equazioni E tale che per ogni n-upla $n_1, ..., n_n$ di numeri naturali, $F(\overline{n}_1, ..., \overline{n}_n) = \overline{m}$ è derivabile da E mediante le regole del calcolo se e solo se $f(n_1, ..., n_n) = m$.[31]

Il lavoro di Gödel sopra ricordato è costituito da note dattiloscritte di una serie di conferenze tenute a Princeton nel febbraio del 1934 e nel 1936 appariva un articolo di Alonzo Church, che si trovava allora a Princeton, dal titolo *An unsolvable problem of elementary number theory* (*Un problema non risolubile della teoria elementare dei numeri*), nel quale veniva individuata la classe delle funzioni λ-convertibili[32] che Church proponeva come corrispettivo rigoroso della classe delle funzioni computabili. Parleremo più avanti degli sviluppi dei λ-sistemi. Qui vogliamo solo ricordare che di questa caratterizzazione si serviva Church nello stesso anno quando pubblicava *A note on Entscheidungsproblem* (*Una nota sul problema della decisione*) ove dava una risposta negativa al problema della decisione per la logica del primo ordine. È questa la prima applicazione della teoria della ricorsività al problema della decisione per una teoria particolare, anche se l'indecibilità di $\mathfrak{P}$ era implicita già nel teorema X della memoria di Gödel del 1931. C'è infatti un nesso essenziale (indagato più tardi in modo sistematico in alcuni lavori di Tarski ed altri) tra completezza e problema della decisione, dato dal fatto, dimostrato da A. Janiczak nel 1953, che se una teoria $\mathfrak{T}$ è assiomatizzata e completa è anche decidibile nel senso che esiste un algoritmo per decidere quando una formula è teorema o no. La dimostrazione è un'applicazione del teorema di Post sopra richiamato, per cui se un insieme e il suo complemento sono ricorsivamente enumerabili, allora l'insieme è ricorsivo. Il

[31] Con F indichiamo un qualsiasi *simbolo* funzionale del linguaggio formale, che diviene così il *nome* della funzione f; per «nome di una funzione» useremo anche il termine «*funtore*».
[32] Il procedimento di Church si fonda su un'idea avanzata già nel 1924 da Moses Schönfinkel in *Über die Bausteine der mathematischen Logik* (*Sugli elementi fondamentali della logica matematica*) e che verrà sviluppata in altro senso negli anni cinquanta in particolare da Haskell B. Curry per la costruzione di una *logica combinatoria* o dei combinatori di cui avremo occasione di riparlare più avanti.

lavoro di Church forniva la prova che la logica del primo ordine è indecidibile e lo stesso valeva – come mostrerà Kleene nel 1936 generalizzando il teorema di incompletezza di Gödel dell'aritmetica di Peano – per molte altre teorie matematiche significative prima fra tutte appunto $\mathfrak{P}$.

Tornando ora alla storia del concetto di funzione computabile, sempre nel 1936 Kleene pubblicava l'articolo *General recursive functions of natural numbers* (*Funzioni generali ricorsive di numeri naturali*) nel quale generalizzava e sistematizzava l'idea di Gödel del calcolo delle equazioni ma dava anche una caratterizzazione per così dire più direttamente «numerica» della classe delle funzioni ricorsive generali, ottenute aggiungendo agli schemi delle pagine 632-33 un ulteriore schema SM fondato sull'impiego dell'operatore μ di minimalizzazione *illimitata*: data una funzione ricorsiva $f(x_1, ..., x_n, x)$ per la quale *esista* almeno un x tale che $f(x_1, ..., x_n, x) = 0$, che obbedisca cioè alla *condizione di regolarità*

$$\forall x_1 \forall x_2, ..., \forall x_n \exists x (f(x_1, ..., x_n, x) = 0),$$

si può definire un'altra funzione *ricorsiva* $g(x_1, ..., x_n)$ ponendo

$$g(x_1, ..., x_n) = \mu x \, f(x_1, ..., x_n, x) = 0$$

ossia il valore di g per ogni n-upla $x_1, ..., x_n$ è dato dal più piccolo x per il quale è verificata la condizione a secondo membro.[33]

Con il concetto di funzione computabile, era possibile caratterizzare in modo preciso anche l'idea intuitiva di insieme effettivamente generabile: la sua controparte rigorosa è il concetto di insieme *ricorsivamente enumerabile* (r.e.): un insieme è r.e. se e solo se coincide col campo di valori di una funzione ricorsiva; gli insiemi decidibili avranno allora come controparte quelli ricorsivi, la cui

[33] L'idea di Kleene si spiega con questa osservazione. Computabili (quindi ricorsive) devono essere tutte le funzioni definibili per induzione. Lo schema di recursione, come mostra l'esempio di Ackermann, afferra solo un aspetto di questo principio generale. Definire una funzione per induzione, significa sostanzialmente definire una relazione il cui carattere funzionale possiamo dimostrare in base al principio di induzione completa. Ora questo principio si può (e lo si vede facilmente) formulare in questi termini: ogni insieme non vuoto di numeri naturali possiede un elemento minimo. È sulla base di questa considerazione che lo schema di definizione mediante l'operatore μ acquista plausibilità: se è soddisfatta la condizione di normalità, allora esiste un minimo nell'insieme degli y tale che $f(x_1, ..., x_n, y) = 0$.

funzione caratteristica cioè è ricorsiva. Su questa base era sviluppabile tutta intera la gamma dei problemi di natura algoritmica.

Sino ad ora abbiamo parlato di funzioni computabili o insiemi decidibili ma non abbiamo direttamente affrontato l'analisi del concetto di algoritmo o di procedura meccanica. È questo concetto che Alan Mathison Turing (1912-54) analizzava nel 1937 nel lavoro *On computable numbers with an application to the Entscheidungsproblem* (*Sui numeri computabili con un'applicazione al problema della decisione*). Al centro dell'attenzione è il procedimento computistico dell'uomo, di cui vengono posti in evidenza gli elementi atomici traducendolo in termini di una macchina «calcolatrice» ideale (la macchina di Turing). Si tratta di una «macchina» che si suppone possedere una «memoria» potenzialmente infinita e un nastro anch'esso potenzialmente infinito; quest'ultimo è pensato diviso in «celle» e la macchina può osservare, in ogni momento, una e una sola di tali celle trovandosi in un determinato «stato interno» (e dispone di un numero finito di tali stati). Essa può compiere quattro operazioni elementari: spostarsi di un passo (di una cella) a destra (D); spostarsi di un passo a sinistra (S); imprimere un simbolo (scelto fra un numero *finito* di simboli base) in una cella vuota; cancellare un simbolo che si trovi in una cella (eventualmente sostituendolo con un altro). Infine la macchina può fermarsi. Si dimostra che con queste operazioni elementari una macchina di Turing riesce a «computare» funzioni nel senso che «atomizza» come sopra le operazioni richieste allo scopo, e a partire dagli argomenti della funzione opportunamente «scritti» sul nastro, si ferma dopo aver stampato il valore della funzione per quegli argomenti. Si può dare forma più astratta a questa impostazione identificando una macchina di Turing con una successione di quadruple ad esempio del tipo seguente

$$q_i \; a_i \; A_k \; q_j$$

da leggersi: trovandosi nello stato interno q_i la macchina osserva il simbolo a_i, compie l'atto A_k e passa nello stato q_j; l'atto A_k può essere: S, D, fermarsi, oppure scrivere il simbolo a_k. (Ovviamente è necessaria una opportuna condizione di «coerenza» per la macchina, tesa a garantire il suo carattere *deterministico*, imponendo che non esistano due quadruple che codificano comandi contraddittori).

È chiaro che da questo punto di vista una funzione sarà computabile se esiste una macchina di Turing che la computa nel senso precedente. Inoltre dovrebbe risultare agevole comprendere che ogni macchina di Turing può essere individuata mediante il suo *numero di Gödel*, che si ottiene semplicemente applicando la tecnica dell'aritmetizzazione all'insieme di quadruple che caratterizza una data macchina. Questo permette a Turing di definire una macchina *universale* $U(x, y)$ che è in grado di simulare il funzionamento di *ogni* singola macchina, nel senso che una volta che le sia dato come input la coppia $<m, n>$, dove m è il numero di Gödel di una qualunque macchina M, il valore $U(m, n)$ calcolato da U coincida col valore $M(n)$ di M applicata a n. Su questa base Turing è in grado di esibire un esempio concreto di problema indecidibile, il problema della fermata (*halting problem*).

Data una macchina M e un input x, la computazione di M si ferma certamente su un output definito o può procedere indefinitamente?

In questi termini il problema riguarda la macchina universale e la sua indecidibilità si può mostrare applicando la tecnica della diagonalizzazione. L'idea è semplicemente la seguente. Supponiamo che esista una macchina $E(x, y)$ in grado di risolvere il problema della fermata: per ogni macchina M, di numero di Gödel m, $E(x, y)$ applicata a m e x, ci darà come output 1 se M si ferma se applicata a x, 0 altrimenti. Definiamo ora una nuova macchina $E'(x)$ tale che per un dato argomento x dia valore 1 se $E(x, x) = 0$ (cioè se la macchina di numero di Gödel x applicata all'argomento x non si ferma), mentre non termina in caso contrario. Sia ora e' il numero di Gödel di questa macchina. Applichiamo la macchina universale U all'argomento $<e', e'>$. Ci sono due casi: o U si ferma oppure no. Nel primo caso, per definizione, avremo $E'(e') = 1$, ma questo avviene se e solo se $E(e', e') = 0$ e questo per definizione di E significa che E' applicata al suo numero di Gödel non si ferma, sicché nemmeno U applicata a $<e', e'>$ può fermarsi: contro l'ipotesi. Ragionamento duale nel secondo caso, cosicché possiamo concludere che non può esistere una macchina di Turing in grado di risolvere il problema della fermata. Se, come Turing sostanzialmente assumeva, ipotizziamo che ogni procedura di decisione è effettuabile da una macchina di Turing, concludiamo che il problema è indecidibile.

Il metodo di approccio di Turing, basato com'è direttamente su un'analisi degli elementi in cui si può scomporre l'attività «reale» di computazione da parte di un soggetto umano più che sulla nozione di computazione (funzioni ricorsive) o di programma (algoritmi), è senz'altro il più convincente e immediatamente afferrabile a livello intuitivo (malgrado le innegabili complessità «meccaniche» che esso comporta). Per questo, più che altre precisazioni, è stato il metodo che ha favorito l'identificazione delle funzioni ricorsive con quelle intuitivamente computabili. Così ricorda Gödel nel 1963 sottolineando come siano stati i lavori di Turing a convincerlo definitivamente della identificabilità tra computabilità e ricorsività generale.

Altre analisi della nozione di algoritmo vennero date in seguito nel 1936 da Emil Post, cui si deve anche, nel 1947, la prima dimostrazione della insolubilità di un problema matematico classico, il problema della parola per semigruppi; da A.A. Markov che nel 1954 introdusse i suoi algoritmi come base per lo sviluppo della sua matematica ricorsiva, ecc.

Ma conviene tornare indietro e considerare alcuni fondamentali lavori di E. Post. Diversamente da Gödel o Turing, Post poneva al centro della sua analisi il concetto di insieme effettivamente generato cui dedicò un lavoro fondamentale, *Recursively enumerable sets of positive integers and their decision problem* (*Insiemi ricorsivamente enumerabili di interi positivi e loro problema della decisione*, 1944), dove viene sviluppato un metodo relativo alla teoria degli insiemi ricorsivamente enumerabili indipendentemente dalla teoria delle funzioni ricorsive e basato sull'uso di particolari sistemi formali introdotti da Post in precedenti lavori, i *sistemi canonici*;[34] l'aggancio con l'approccio «funzionale» era dato dai teoremi –

[34] All'origine di questo approccio sta un'analisi del concetto generale di sistema formale come procedura sistematica per generare oggetti complessi, i teoremi, che sono *parole* formulate entro un dato alfabeto. In questo senso quello di Post è un tentativo più «logico» che matematico in un doppio senso: da una parte costituisce un ripensamento dell'idea di sistema formale (Hilbert), dall'altra un tentativo estremamente audace di afferrare direttamente il concetto base della matematica costruttiva: il concetto di insieme potenzialmente infinito generato per approssimazioni successive (si pensi a Gentzen). Le potenzialità del tentativo di Post sono emerse in modo netto nel 1961, quando Raymond Smullyan in *Theory of formal systems* (*Teoria dei sistemi formali*) si è servito di questi metodi per una presentazione estremamente elegante della teoria della ricorsività.

dimostrati da Post – che in forma intuitiva abbiamo già dato e cioè: la funzione caratteristica f_M di un insieme M è computabile se e solo se M e il suo complemento $\overline{M}$ sono entrambi ricorsivamente enumerabili e una funzione f è computabile se e solo se l'insieme delle coppie $<x, f(x)>$ è ricorsivamente enumerabile.

L'osservazione chiave per capire l'interesse della proposta di Post è che ogni sistema formale nel senso hilbertiano (ogni logica simbolica, usando il termine di Post) utilizzando un linguaggio finitamente descritto e assiomi e regole decidibili, dà origine a insiemi di teoremi che sono appunto ricorsivamente enumerabili (passando, via aritmetizzazione, agli insiemi numerici corrispondenti). Questo vale in particolare per i teoremi dell'aritmetica di Peano, per le leggi logiche (una volta che si accetti l'identificazione – non costruttivamente provata, come sottolineava Church – tra dimostrabile e valido garantitaci dal teorema di completezza di Gödel) ecc. Si può dire che la nozione di sistema canonico di Post era essenzialmente motivata proprio dal desiderio di catturare astrattamente quella caratteristica di generabilità mediante schemi finiti che distingue i teoremi di ogni sistema formale concreto.

Domanda naturale: esistono insiemi ricorsivamente enumerabili ma non ricorsivi? La risposta viene considerando i diversi problemi indecidibili che già abbiamo incontrato. Il primo esempio di insieme ricorsivamente enumerabile ma non ricorsivo ci è dato dalla famiglia delle leggi logiche di un linguaggio del primo ordine sufficientemente ricco, in forza del teorema di Church; un altro esempio, come Kleene e Post fecero vedere, dall'insieme di teoremi dell'aritmetica di Peano (questo per il teorema di incompletezza di Gödel in vista di quel collegamento fra completezza e decidibilità per teorie assiomatizzate stabilito in generale dal già ricordato teorema di Janiczak). Altri esempi di insiemi r.e. non ricorsivi si possono ottenere sfruttando l'esistenza di una macchina di Turing universale o anche, come fece vedere Post, più semplicemente sfruttando la enumerabilità ricorsiva dei sistemi canonici. È sufficiente disporre in una sorta di matrice infinita gli insiemi r.e. (via l'enumerazione dei sistemi canonici oppure ricorrendo alla macchina di Turing universale) e procedendo per diagonalizzazione costruire un insieme il cui complemento non è r.e. e che quindi non è ricorsivo.

Nell'articolo citato, in cui prevedeva per la teoria delle funzioni

ricorsive un ruolo nella matematica combinatoria secondo solo a quello dello stesso concetto di numero naturale, Post analizzava in generale il problema di classificare gli insiemi r.e. non ricorsivi in termini della loro complessità. Un ordinamento della complessità per Post è dato dalla possibilità di ricondurre il problema della decisione per un insieme a quella di un altro e la domanda diviene allora quella di vedere che tipi di riduzioni possano esistere. Post introduceva quattro tipi di riducibilità: la riducibilità uno-uno, la riducibilità molti-uno, la riducibilità via tavole di verità limitate e illimitate. In tutti questi casi Post indaga se esistono insiemi r.e. *universali*, nel senso che, considerando le singole nozioni di riduzione, ogni insieme r.e. non ricorsivo si riduce all'universale: gli insiemi universali sono cioè quelli a cui è possibile ricondurre il problema della decisione di ogni altro insieme. Nel corso di queste analisi Post individuava diversi tipi di insiemi r.e.: gli insiemi creativi, gli insiemi proiettivi, immuni, semplici, ipersemplici, completi ecc., in base alla loro possibilità di essere candidati al ruolo di insiemi universali. Nei diversi casi Post analizzava se esistevano insiemi r.e. non riducibili fra di loro e poteva provare ad esempio che gli insiemi semplici hanno un tipo di indecidibilità meno complesso di quelli completi, rispetto all'ordine dato dalla riducibilità via tavole limitate, o che, ancora, gli insiemi creativi non sono riducibili ad insiemi ipersemplici via tavole di verità illimitate. Questo significava in altre parole che esistono insiemi r.e. che non hanno lo stesso «grado» di indecidibilità nel senso tecnico preciso che non sono uno riducibile all'altro secondo le varie accezioni di riducibilità sopra elencate.

Per ultima Post considerava la nozione di riducibilità più generale, che era stata introdotta da Turing nel 1939 in *Systems of Logic based on ordinals* (*Sistemi di logica basati sugli ordinali*) e che si fondava sul concetto di ricorsività relativa: un insieme A di numeri naturali si dice *ricorsivo in* un insieme B quando la funzione caratteristica f_A del primo risulta computabile con schemi ricorsivi utilizzando, ogni qualvolta sia necessario alla computazione di f_A, dei valori che f_B assume su dati argomenti. In altre parole, come scriveva Turing, computiamo f_A immaginando di avere un *oracolo* che ci dà i valori richiesti di f_B.

La nozione di ricorsività relativa induce un preordine $<$ le cui classi di equivalenza sono note oggi come *gradi di insolubilità* in

stretta analogia a quanto detto sopra sulle altre nozioni di riducibilità. È facile convincersi che quest'ultima nozione di riducibilità è la più generale di tutte, in quanto non pone vincoli sul tipo di processo mediante il quale passiamo dal problema di decisione di un insieme a quello di un altro. Il problema che Post poneva concludendo il suo articolo – noto oggi *tout court* come *problema di Post* – era se esistono insiemi r.e. non ricorsivi che non hanno lo stesso grado, vale a dire uno dei quali non sia ricorsivo nell'altro. Esso non tardò a rivelarsi estremamente complesso. Le tecniche fino ad allora utilizzate per costruire insiemi non ricorsivi non erano sufficienti per dare una risposta e sarà solo nel 1956, come accenneremo più avanti, che il problema verrà risolto mostrando come lo studio dei gradi presenti aspetti estremamente significativi dal punto di vista matematico.

Un problema generale rimaneva però sullo sfondo: entro che misura tutte queste analisi del concetto di computabile, di decidibile, di effettivamente generabile sono adeguate e che rapporti hanno fra di loro?

Un risultato fondamentale cui pervennero vari autori – ad esempio Kleene, Church, Turing, Rosser – stabiliva che le varie definizioni – in termini di funzioni ricorsive, o di funzioni «rappresentabili» in un dato sistema (Gödel e Church), o di macchina di Turing, o infine di sistemi canonici (Post) – sono tutte *equivalenti*, ossia individuano la stessa classe di funzioni. È questo uno degli argomenti fondamentali che deponevano a favore della cosiddetta *tesi di Church* (che questo autore avanzò esplicitamente nel primo degli articoli del 1936 sopra citati) secondo la quale, accettato come è ovvio che ogni funzione ricorsiva è computabile, vale anche l'inverso ossia *ogni funzione* (intuitivamente) *effettivamente computabile è ricorsiva*: in altri termini, le funzioni computabili sono tutte e sole le funzioni ricorsive.[35]

[35] Si noti che l'affermazione della coincidenza della classe delle funzioni ricorsive con la classe delle funzioni effettivamente computabili viene chiamata *tesi* o *ipotesi* (e non, ad esempio, *teorema*) di Church, perché si tratta di una congettura che afferma la coincidenza (estensionale) di due concetti, l'uno dei quali (quello di funzione ricorsiva) precisato rigorosamente a livello formale, l'altro (quello di funzione effettivamente computabile) dato solo intuitivamente e in accezione assai vaga. La tesi di Church non è quindi suscettibile di dimostrazione rigorosa: si possono solo portare argomenti pro o contro di essa. La tesi è sostanzialmente accettata dalla maggioranza dei logici e dei matematici anche se nel 1959 sono state

La dimostrata stabilità del concetto di computabilità dava senso allo sviluppo di una vera e propria *teoria della ricorsività* su basi sistematiche, quale fu intrapresa da S.C. Kleene in una serie di lavori a partire dal 1936. Kleene adottava sistematicamente la tecnica della gödelizzazione riferendosi esplicitamente (ma talora solo implicitamente) ad uno specifico sistema formale per definire le funzioni ricorsive: il calcolo funzionale di Herbrand-Gödel. Sfruttando queste tecniche, Kleene era in grado di provare alcuni risultati che offrono una base compatta a tutta la teoria. Per cominciare Kleene mostrava come sia possibile definire un predicato ricorsivo *primitivo* T^{n+2} a $n+2$ posti per ogni $n \geqslant 0$, che traduce aritmeticamente la relazione che sussiste tra una funzione di indice e, argomenti $x_1, \ldots, x_n$ ed y, quando y è una computazione, sulla base del formalismo adottato, del valore che la funzione assume per gli argomenti $x, \ldots, x_n$. In questo modo essendo computazioni, funzioni e possibili argomenti sullo stesso piano, in quanto tutti codificati in termini di numeri naturali, i fatti centrali riguardanti le funzioni ricorsive e gli insiemi r.e. si possono esprimere in termini del predicato T^{n+2} al variare di n. Da questo scende facilmente come corollario un teorema di forma normale che ci permette di esprimere il valore di ogni funzione ricorsiva generale in forma canonica. Supponiamo infatti di applicare l'operatore di minimalizzazione μ al predicato T^{n+2} e di ottenere $\mu y T^{n+2}(e, x_1, \ldots, x_n, y)$, ossia il numero assegnato alla più piccola computazione della funzione di indice e applicata agli argomenti $x_1, \ldots, x_n$. Kleene mostrava che è possibile definire una funzione U ricorsiva primitiva che «estrae» da ogni computazione il valore che essa dà a determinati argomenti, cosicché per ogni funzione ricorsiva $f(x_1, \ldots, x_n)$ vale l'equazione

$$(*) \qquad f(x_1, \ldots, x_n) = U(\mu y T^{n+2}(e, x_1, \ldots, x_n, y))$$

sollevate precise obiezioni contro di essa in particolare da parte di Rosza Peter (che riteneva la classe delle funzioni ricorsive troppo ampia per essere adeguata) e da parte di Laszlo Kalmar (che, viceversa, riteneva tale classe troppo ristretta per comprendere tutte le funzioni effettivamente calcolabili). A queste obiezioni ha risposto fra gli altri Elliott Mendelson (1963). Lo sviluppo più interessante a questo proposito è tuttavia un altro e si basa non sul confronto tra la nozione di funzione ricorsiva con quella di funzione computabile, ma piuttosto con quella di funzione costruttiva nel senso intuizionista. Lavori in questo senso sono dovuti a Myhill, Mendelson, Kreisel e altri.

ossia la funzione, applicata alla minima computazione come sopra detto, fornisce il valore di tale computazione. L'espressione precedente è una formulazione del cosiddetto teorema di *forma normale* di Kleene, secondo il quale *ogni* funzione ricorsiva può essere espressa nella forma (*) per un opportuno *e*.

Non è difficile riconoscere nel predicato T^{n+2} la controparte aritmetica della macchina universale di Turing e la cosa ha un'importanza fondamentale non solo in quanto ci dà un metodo universale per computare ogni funzione ricorsiva, ma anche perché mette in rilievo due fatti: 1) il sostanziale riferimento non costruttivo a una totalità infinita (quella dei numeri naturali) implicito in tutte le caratterizzazioni da noi presentate per le funzioni ricorsive; ciò è messo in evidenza dal ricorso all'operatore di minimalizzazione illimitata nel teorema di forma normale, che sostanzialmente significa che per caratterizzare la computabilità è *necessario* un riferimento non costruttivo di tipo esistenziale alla totalità delle computazioni.[36] 2) Il fatto che la proprietà di una funzione di essere ricorsiva non comporta di necessità l'essere definita per *ogni* argomento, ma piuttosto l'essere computabile su ogni argomento per il quale è definita. Anche in questo caso è l'operatore μ che pone in luce questa situazione. Infatti l'applicazione di questo operatore ci porta ad un valore solo nel caso che sia soddisfatta la condizione di *regolarità* vista sopra.

La natura stessa del formalismo introdotto da Kleene porta così naturalmente a privilegiare il concetto di *funzione parziale ricorsiva* (f.p.r.) nella cui definizione l'operatore μ può essere applicato anche a predicati non necessariamente regolari. Saranno così funzioni *ricorsive generali* quelle funzioni ricorsive parziali che sono defini-

[36] Questo ineliminabile ricorso a un quantificatore esistenziale cui non è possibile assegnare un significato costruttivo senza già presupporre quella precisazione del concetto di funzione computabile che tramite esso si intende chiarire, ha portato gli intuizionisti, in particolare Heyting nel 1960, a rigettare l'identificazione proposta dalla tesi di Church, nel senso che gli intuizionisti ritengono quello di *funzione computabile* un concetto primitivo e quindi rifiutano come circolare ogni tentativo di una sua precisazione rigorosa. Ciò ribadisce ancora una volta che la teoria della ricorsività costituisce un tentativo di precisazione del concetto di effettivo all'*interno della matematica classica* ossia, se ci si permette il gioco di parole, non pretende di presentarsi come una chiarificazione *effettiva* dell'*effettivo*.

te per ogni argomento. Il passaggio dalle funzioni generali a quelle parziali permette a Kleene di provare il cosiddetto *teorema di enumerazione* in base al quale, sfruttando il predicato T di sopra, è possibile ottenere un'enumerazione effettiva di tutte le funzioni ricorsive parziali. Il fatto contrasta con quanto visto sulle funzioni primitive ricorsive: allora, diagonalizzando, avevamo provato che l'enumerazione delle ricorsive primitive non poteva essere data da una funzione a sua volta ricorsiva primitiva; considerando le funzioni parziali sfuggiamo alla diagonalizzazione proprio in quanto la funzione che le enumera non si limita alle totali. È il teorema di enumerazione che ci dà la misura di come sia più fondamentale il concetto di f.p.r. rispetto a quello di funzione ricorsiva generale e ciò ha un immediato corollario sul piano dei problemi di decisione: è possibile determinare in modo effettivo se una funzione ricorsiva è parziale o totale? Orbene la risposta è *negativa* (ossia il problema non è *decidibile*) ed è questo un ulteriore risultato centrale che proviamo con la diagonalizzazione come l'analogo problema della fermata per le macchine di Turing.

Questo risultato forniva un ulteriore esempio del fatto che esistono insiemi *ricorsivamente enumerabili* ma non *ricorsivi*, e dava senso al tipo d'approccio sviluppato da Post e alle sue indagini circa la riducibilità degli insiemi cui abbiamo prima accennato. È in questo contesto che Kleene generalizzava ed unificava i precedenti risultati di indecidibilità di Church e di Gödel, ponendo in evidenza che il teorema di indecidibilità di Gödel, dal punto di vista della problematica della ricorsività si può formulare dicendo che l'insieme dei teoremi dell'aritmetica di Peano al primo ordine non è ricorsivo, in quanto non è r.e. il suo complemento.

Il passaggio dai risultati di indecidibilità alla esistenza di insiemi ricorsivamente enumerabili non ricorsivi è reso esplicito da un importante collegamento che si può stabilire tra forma dei predicati ed insiemi che determinano: un predicato ricorsivo n-ario R determina come sappiamo un insieme ricorsivo; un predicato ricorsivo R cui sia applicato un quantificatore esistenziale illimitato, e quindi della forma $\exists yR$, determina un insieme ricorsivamente enumerabile. Un insieme sarà allora ricorsivamente enumerabile ma non ricorsivo quando sarà definito facendo ricorso *in modo essenziale* ad una quantificazione esistenziale: è questo il caso appun-

to dell'insieme degli indici delle funzioni ricorsive *totali*, come si può vedere considerando il predicato di Kleene. In questo contesto il teorema di Post, sopra ricordato, ammette la seguente formulazione: un insieme è ricorsivo se e solo se si può definire mediante la quantificazione esistenziale o la quantificazione universale di un predicato ricorsivo. Viene naturale allora una disposizione di questo tipo per i predicati:

$$R(x_1, \ldots, x_n) \begin{array}{l} \nearrow \exists x R(x_1, \ldots, x_n, x) \\ \searrow \forall x R(x_1, \ldots, x_n, x) \end{array}$$

dove R è genericamente un predicato ricorsivo a n posti ($n \geqslant 1$, finito) e le frecce indicano inclusioni proprie. In questo modo si può «misurare» il grado di effettività di un predicato (di un insieme, di una funzione) e sorge spontanea l'idea di proseguire questa descrizione considerando prefissi via via più complessi ossia sempre meno «effettivi» (vale a dire con più numerosi riferimenti alla totalità dei numeri naturali). Come mostrava un ulteriore, importantissimo, risultato di Kleene (1943) la cosa è possibile, ossia si può stabilire una gerarchia (detta *gerarchia aritmetica*) ottenuta indipendentemente negli stessi anni da Mostowski:

$$R(x_1, \ldots, x_n) \begin{array}{l} \nearrow \exists x R(x_1, \ldots, x_n, x) \to \forall y \exists x R(x_1, \ldots, x_n, x, y) \to \exists \forall \exists R \ldots \\ \searrow \forall x R(x_1, \ldots, x_n, x) \to \exists y \forall x R(x_1, \ldots, x_n, x, y) \to \forall \exists \forall R \ldots \end{array}$$

dove R è un attributo numerico ricorsivo. «Si può stabilire» va inteso nel senso che ad ogni passo della gerarchia si individuano attributi non riducibili a nessuno di quelli ottenuti in passi precedenti; e quindi l'estensione non è banale. Si usa indicare con notazione unitaria col simbolo Π_n^0 ($n > 0$) la classe degli insiemi descritti da predicati il cui prefisso inizia con un quantificatore universale e contiene $n - 1$ alternanze di quantificatori; ad esempio Π_3^0 sarà la classe degli insiemi esprimibili da predicati della forma $\forall \exists \forall\ R$ (si noti che più quantificatori dello stesso tipo immediatamente susseguentisi l'un l'altro si possono *contrarre* in un unico quantificatore di quel tipo: così ad esempio $\forall\forall\forall\exists$ si può contrarre in $\forall\exists$). Il prefisso Π_0^0 (come l'analogo Σ_0^0 che ora introdurremo) in-

dica assenza di quantificatori. Viceversa si indica con Σ_n^0 l'analoga classe degli insiemi esprimibili da predicati il cui prefisso inizia con un esistenziale. In questa notazione compatta la gerarchia assume quindi la seguente forma

$$\Pi_0^0 = \Sigma_0^0 \begin{array}{l} \nearrow \ \Pi_1^0 \rightarrow \Pi_2^0 \rightarrow \Pi_3^0 \rightarrow \ldots \\ \qquad \times \quad \times \\ \searrow \ \Sigma_1^0 \rightarrow \Sigma_2^0 \rightarrow \Sigma_3^0 \rightarrow \ldots \end{array}$$

dove le frecce indicano l'inclusione propria. Tale gerarchia gode di notevoli proprietà. Intanto, come già sopra accennato, essa non è banale, vale a dire, in conformità al teorema di gerarchia di Kleene (1943), per ogni n esiste un insieme A tale che $A \in \Pi_{n+1}^0$ ($\in \Sigma_{n+1}^0$) ma $A \notin \Pi_n^0$ né $A \in \Sigma_{n+1}^0$ ($A \notin \Sigma_n^0$ né a Π_{n+1}^0). Ancora, in generale la riunione $\Pi_n^0 \cup \Sigma_n^0$ è contenuta in modo proprio nella intersezione $\Pi_{n+1}^0 \cap \Sigma_{n+1}^0$; l'unica eccezione è data da $\Delta_0^0 = \Pi_0^0 = \Sigma_0^0 = \Pi_1^0 \cap \Sigma_1^0$ il che è un modo appunto di esprimere il fatto che un insieme A è ricorsivo quando tanto A quanto il suo complemento $\overline{A}$ sono ricorsivamente enumerabili. Ultimo risultato significato in questo contesto è la generalizzazione del teorema di enumerazione di Kleene (1943) che dà una sorta di rappresentazione unitaria per tutti gli insiemi della gerarchia aritmetica ricorrendo al predicato T sopra definito. Risultati analoghi, come dimostrarono Kleene e Post, valgono anche per la nozione di ricorsività relativa.

Sono questi alcuni dei temi centrali della ricerca sul concetto di effettivo fino agli anni quaranta; conviene però, per compattezza di discorso, accennare già ora ad alcuni ulteriori sviluppi della teoria della ricorsività anche se cronologicamente posteriori al periodo qui considerato. Gli sviluppi di cui parliamo si concentrano su un problema comune: quello di saggiare la possibilità di ulteriori classificazioni per gli attributi (predicati, funzioni, insiemi) lungo le linee indicate rispettivamente da Kleene (con la costruzione della gerarchia aritmetica) e da Post (con i gradi di irresolubilità). Per quanto riguarda il primo aspetto il passo naturale fu di estendere la gerarchia aritmetica considerando la classe degli attributi numerici definibili tramite predicati che ammettessero la quantificazione illimitata non solo (come nel caso della gerarchia aritmetica) sulle variabili individuali, ma anche sulle variabili

funzionali (o per insiemi; passando cioè a un linguaggio del secondo ordine). Anche in questo caso la classificazione in base alla forma della definizione linguistica riflette una «misura» del grado di effettività: la classe degli attributi così ottenuti viene detta degli *attributi analitici* e Kleene in un articolo del 1955 ha mostrato che anche in questo caso può ottenersi una *gerarchia* (detta appunto *analitica*) nel senso che tutti questi attributi si possono porre in una delle forme

$$A(x_1, \ldots, x_n) \begin{array}{l} \nearrow \exists f \forall x\, R \rightarrow \forall g \exists f \forall x\, R \ldots \\ \searrow \forall f \exists x\, R \rightarrow \exists g \forall f \exists x\, R \ldots \end{array}$$

o, in forma abbreviata

$$\Delta_0^1 \begin{array}{l} \nearrow \Sigma_1^1 \rightarrow \Sigma_2^1 \ldots \\ \searrow \Pi_1^1 \rightarrow \Pi_2^1 \ldots \end{array}$$

dove A è un predicato (attributo) aritmetico e R è ricorsivo. Questa gerarchia i cui elementi vengono indicati con Σ_n^1, rispettivamente Π_n^1, gode di molte delle proprietà viste per la gerarchia aritmetica ma con una importante eccezione: $\Delta_0^1 \neq \Sigma_1^1 \cap \Pi_1^1$, vale a dire gli attributi numerici *non* sono la parte comune agli analitici del primo gradino della gerarchia. Gli elementi di tale intersezione coincidono con quegli *attributi iperaritmetici* che nello stesso articolo Kleene dimostrava essere individuabili mediante una gerarchia indiciata da ordinali costruttivi che prolungava quella aritmetica. Essi possono intuitivamente essere considerati come una sorta di «limite» cui tendono gli aritmetici nel passaggio agli analitici. La teoria della ricorsività veniva così a ricollegarsi direttamente con quelle indagini che, a partire dalla scuola degli analisti francesi, erano state compiute nel tentativo di colmare l'abisso tra il «dato immediatamente», il «manipolabile», l'«effettivo», e il transfinito cantoriano nello studio degli insiemi di numeri reali tanto dal punto di vista topologico quanto da quello di teoria della misura. Come già sottolineava Mostowski nei suoi primi lavori sull'argomento, le gerarchie aritmetica e analitica presentano infatti notevoli analogie (e differenze significative) con quelle degli insiemi

proiettivi e analitici studiate da Borel e Lebesgue e quindi sviluppate da Michel Suslin e Nicolai Lusin, e più in generale con alcuni temi centrali della teoria descrittiva degli insiemi.

Sulla base dei risultati di Kleene e Mostowski, negli anni cinquanta John Addison poteva anzi mostrare come più che di un'analogia si trattasse di un'unica teoria, la teoria delle gerarchie, vista da diverse prospettive e come la teoria descrittiva degli insiemi si potesse ottenere relativizzando la teoria delle gerarchie della ricorsività usando il seguente vocabolario di traduzione:

Funzione ricorsiva	Funzione continua
Insieme r.e.	Insieme aperto
Insieme iperaritmetico	Insieme boreliano
Σ_1^1-insieme	Insieme analitico
Insieme nella gerarchia analitica	Insieme proiettivo.

Sarà negli anni sessanta che con lo svilupparsi delle generalizzazioni della teoria della ricorsività si vedranno meglio i rapporti tra la teoria delle gerarchie, i problemi di definibilità e l'uso di estensioni dei linguaggi elementari. Quello che salta agli occhi è che il collegamento fra funzioni ricorsive e funzioni continue pone in luce la stretta affinità – percepita fin dall'inizio da matematici come Borel e Baire, ma soprattutto da Brouwer – tra costruttività e continuità.

Se le gerarchie aritmetica e analitica si fondano essenzialmente sulla *forma* delle definizioni (i blocchi di quantificatori), la nozione di ricorsività relativa si presta ad un altro tipo di gerarchizzazione sulla base dei già nominati *gradi* (*di insolubilità*) ossia le classi di equivalenza rispetto alla Turing-riducibilità. Nel 1944 Post mostrava che, assumendo come relazione d'ordine la $\leqslant$ indotta dalla riducibilità i gradi possono essere ordinati formando una struttura che si riconosce essere un semireticolo superiore nel senso che dati due qualunque gradi α e β esiste sempre un minimo grado γ tale che $\alpha \leqslant \gamma$ e $\beta \leqslant \gamma$. Come hanno dimostrato Kleene e Post nel 1954 non vale la stessa proprietà «verso il basso» ossia non esiste in generale, dati due gradi α e β, un massimo grado γ per il quale $\alpha \geqslant \gamma$ e $\beta \geqslant \gamma$. I gradi formano un insieme di cardinalità uguale a quella del continuo ($\aleph_1$, accettando l'ipotesi del continuo) e

nel 1960 Schoenfield ha dimostrato che esiste un insieme di gradi che ha cardinalità $\aleph_1$ e i cui elementi sono fra loro inconfrontabili.

Il grado più basso, indicato con 0, è evidentemente quello delle funzioni ricorsive. Se con W_x indichiamo l'x-esimo r.e. e con K l'insieme $\{x \mid x \in W_x\}$ avremo che il grado di K, di solito indicato con $0'$, è strettamente maggiore di 0. Relativizzando la definizione di K, si ottiene un'operazione $'$ sui gradi, detta *jump*, per cui per ogni grado α $\alpha < \alpha'$. Esistono quindi gradi arbitrariamente alti, e inoltre, come dimostrato da Sacks, l'insieme dei gradi r.e. è denso. Fatto fondamentale è che esiste un a tale che $0 < a < 0'$ e questo fornisce una soluzione al problema di Post, come fu data nel 1956 da Richard M. Friedberg in *Two recursively enumerable sets of incomparable degrees of unsolvability* (*solution of Post's problem* 1944) ([*Due insiemi ricorsivamente enumerabili con gradi di irresolubilità non confrontabili* (*soluzione del problema di Post* 1944)] e da A.A. Mŭcnik in *Risposta negativa al problema di riducibilità degli algoritmi*: entrambi costruivano due insiemi nessuno dei quali è ricorsivo nell'altro introducendo il cosiddetto *metodo di priorità* che si è rivelato estremamente utile in successive ricerche e che come vedremo costituisce una sorta di antenato del metodo del *forcing* di Cohen.

Visto il ruolo centrale delle nozioni legate al concetto di effettivo nell'ambito della metamatematica finitista, non stupisce che numerose ricerche – dopo la creazione della teoria della ricorsività – abbiano avuto come tema il collegamento fra insiemi r.e. e sistemi formali. Del resto abbiamo visto sopra come l'idea di considerare rapporti di riducibilità tra i problemi di decisione per insiemi diversi che è all'origine dell'indagine sui gradi, partisse in Post proprio dall'analisi di questi sistemi. La classificazione degli insiemi r.e. ebbe così per converso un impatto significativo nello studio del problema della decisione per teorie, fornendo definizioni precise per nozioni spesso abbastanza indeterminate, prima fra tutte quella di *assiomatizzabilità*, di cui abbiamo avuto occasione di parlare più volte. Sempre supponendo che le nostre teorie siano definite su un linguaggio effettivamente specificabile (le cui costanti extralogiche siano così al più numerabili) sappiamo che è possibile introdurre una distinzione fondamentale all'interno delle teorie che si basano sulla logica del primo o del secondo ordine: diciamo

finitamente assiomatizzate le teorie il cui insieme di assiomi extralogici è finito, mentre *assiomatizzata* o *ricorsivamente assiomatizzata* è ogni teoria il cui insieme di assiomi sia ricorsivo. *Assiomatizzabili* (o *finitamente assiomatizzabili*) sono quelle teorie che sono equivalenti – nel senso che hanno gli stessi teoremi – a teorie assiomatizzate o finitamente assiomatizzate. Esempi di teorie finitamente assiomatizzabili sono la teoria dei gruppi o quella degli anelli, la teoria hilbertiana delle *Grundlagen* e l'aritmetica di Peano al secondo ordine, la teoria degli insiemi nella formulazione $\mathfrak{GB}$ di Gödel, ecc. Esempi di teorie non finitamente assiomatizzabili sono l'aritmetica di Peano al primo ordine, la teoria dei numeri reali, sempre al primo ordine, il sistema $\mathfrak{ZF}$, ecc.

Un fatto che si osserva immediatamente è che il passaggio dal primo al secondo ordine permette spesso di ottenere l'assiomatizzabilità finita; è questo ad esempio il caso dell'aritmetica di Peano e della teoria dei numeri reali. Un problema fondamentale è quindi quello di sapere quando una teoria è finitamente assiomatizzabile al primo ordine. Ci sono diverse tecniche per ottenere questo, come vedremo anche più avanti, e alcune si basano sull'uso della teoria della ricorsività. È così ad esempio che nel 1952 Ryll-Nardzewski ha dimostrato che l'aritmetica di Peano al primo ordine *non* è finitamente assiomatizzabile e alla fine degli anni cinquanta M. Rabin ha riottenuto lo stesso teorema in maniera più semplice. Con tecniche che ancora fanno riferimento alla teoria della ricorsività R. Montague nel 1957 ha indagato il rapporto fra la teoria di Zermelo senza assioma di rimpiazzamento e $\mathfrak{ZF}$, mostrando che non si tratta di un'estensione ottenibile con un numero finito di assiomi.

Dal punto di vista del problema della decisione l'interesse dell'assiomatizzabilità sta – come aveva mostrato Post – nel fatto che sia usando la logica del primo ordine che quella del secondo, tutte le teorie assiomatizzabili hanno insiemi di teoremi r.e., e questo in forza della natura decidibile degli assiomi logici e delle regole del calcolo. Il fatto che una teoria sia decidibile se completa e assiomatizzata come espresso in modo generale da Janiczack nel teorema già ricordato, non è che la controparte per teorie del teorema di Post secondo il quale un insieme è ricorsivo quando sia esso che il suo complemento sono r.e. La difficoltà, nel caso di teorie come l'aritmetica di Peano, sta allora nel fatto che l'insieme dei «non

teoremi» non è r.e. La non dominabilità di questo insieme è collegata alla estrema non effettività dell'insieme delle formule vere della struttura dei naturali che è addirittura iperaritmetico, come mostra il teorema di Tarski sulla non definibilità della verità nell'aritmetica. Ma altri aspetti del problema furono indagati negli anni cinquanta. Uno dei risultati più interessanti (notevole anche per le conseguenze che ha avuto nella filosofia della scienza) è il teorema di W. Craig secondo il quale ogni teoria il cui insieme di teoremi sia r.e. è *sempre* assiomatizzabile cosicché l'assiomatizzabilità diventa una proprietà ricorsiva indipendente dalla scelta degli assiomi. Ancora Craig, in collaborazione con Kleene, negli stessi anni affrontava il problema dell'assiomatizzabilità mediante aggiunta di nuovi predicati, problema interessante anche dal punto di vista generale in quanto permette di affrontare formalmente l'analisi del ruolo dei termini teorici nella derivazione di risultati empirici nelle scienze naturali. Altre caratterizzazioni significative in questo contesto riguardano l'assiomatizzabilità *indipendente*, che si ha quando ogni assioma non è dimostrabile a partire dagli altri. Come Kreisel provava nel 1957, ad esempio, esiste in questo caso un legame significativo con gli insiemi ipersemplici di Post.

Il concetto più significativo però per quanto riguarda la nozione di decidibilità è quello di teoria *essenzialmente indecidibile* che fu introdotto da Tarski nel 1949 e che trova il suo esempio più cospicuo proprio nell'aritmetica di Peano al primo ordine: una teoria è essenzialmente indecidibile quando l'insieme V dei teoremi e l'insieme F delle formule refutabili (le formule cioè la cui negazione è teorema) risultano *inseparabili* con un insieme ricorsivo X tale cioè che

$$V \subseteqq X \text{ e } F \subseteqq \overline{X}.$$

Il concetto si può esprimere in altri termini dicendo che $\mathfrak{T}$ è essenzialmente indecidibile quando ogni sua estensione nello stesso linguaggio che sia coerente è indecidibile.

Gli anni sessanta hanno visto fiorire ricerche sulla classificazione delle teorie indecidibili in termini del tipo di non effettività del loro insieme di teoremi e si è dimostrato ad esempio che esistono teorie i cui insiemi di teoremi appartengono ad ogni grado di insolubilità. Ma la conclusione generale è stata che nella mag-

gior parte dei casi le teorie assiomatizzate più note sono sostanzialmente equivalenti dal punto di vista del problema della decisione e che un'analisi più fine delle loro differenze si può condurre solo abbandonando il punto di vista estensionale della teoria della ricorsività per la considerazione più diretta di tipo intensionale della loro struttura dimostrativa. Come affermava Kreisel nel 1971 «la teoria della dimostrazione inizia dove finisce la teoria della ricorsività».

Torneremo più avanti su esempi specifici e su altre tecniche connesse al problema della decisione, ora vogliamo affrontare un altro aspetto della questione, cioè lo studio dei metodi per dimostrare l'*indecidibilità* di una teoria. Come le ricerche degli anni trenta mostravano, le vie erano sostanzialmente due: o un modo diretto, come quello utilizzato da Gödel, o una strategia indiretta che riconduce il problema della decisione di una teoria $\mathfrak{T}'$ di cui ci è già nota l'indecidibilità al problema della decisione per la teoria $\mathfrak{T}$ in esame, cosicché se $\mathfrak{T}$ fosse decidibile lo sarebbe, contro l'ipotesi, anche $\mathfrak{T}'$. Non ci fermeremo qui su questo secondo tipo di approccio che ha avuto diversi sviluppi ma che in ultima istanza trova il suo pilastro nel concetto di interpretazione di una teoria in un'altra di cui avremo occasione di parlare più avanti e nel ricorso a teorie essenzialmente indecidibili; ci basti osservare che in questo contesto divengono rilevanti tutte le considerazioni sui diversi tipi di riducibilità analizzati da Post.

Vogliamo soffermarci invece su una questione di portata più generale e più strettamente legata alla teoria della ricorsività. Come si può dimostrare in modo *diretto* che una teoria è indecidibile? La tecnica usata da Gödel si basava sul duplice fatto che:

1) le relazioni metamatematiche fondamentali riguardanti la teoria in esame sono ricorsive (primitive)

2) relazioni e funzioni ricorsive sono rappresentabili all'interno della teoria.

Quello che la ricerca degli anni cinquanta ha posto in luce è la centralità del concetto di rappresentabilità concetto il cui ruolo non si limita come vedremo al solo problema della indecidibilità delle teorie. Nel suo articolo Gödel parlava di proprietà numeriche *entscheidungsdefinit* ed il concetto corrisponde alla nozione di

rappresentabilità formale da noi data nella nota 7 a pag. 569 e che viene anche detta rappresentabilità *forte*. Si parla invece di rappresentabilità *debole* di una relazione numerica R all'interno di una teoria $\mathfrak{T}$ formulata nel linguaggio dell'aritmetica quando esiste una formula $\mathscr{A}$ con n variabili libere per cui

$$\text{per ogni sequenza di naturali } a_1, \ldots, a_n$$
$$\text{si ha } \langle a_1, \ldots, a_n \rangle \in R \text{ sse } \vdash_{\mathfrak{T}} \mathscr{A}(\bar{a}_1, \ldots, \bar{a}_n)$$

dove $\bar{a}_i$ indica il termine che nella teoria $\mathfrak{T}$ denota il numero a_i.

Riferendosi a questa forma debole di rappresentabilità possiamo riassumere i risultati centrali ottenuti da diversi ricercatori (P. Bernays, H. Putnam, A. Tarski, A. Mostowski e R. Robinson) dicendo che una teoria assiomatizzabile $\mathfrak{T}$ risulta indecidibile se

1) *ogni* insieme ricorsivo è debolmente rappresentabile in $\mathfrak{T}$ oppure

2) *qualche* insieme r.e. ma non ricorsivo è debolmente rappresentabile in $\mathfrak{T}$.

Le due condizioni forniscono metodi distinti per provare l'indecidibilità di una teoria, anche se nel secondo caso molti dei sistemi noti in cui è rappresentabile qualche insieme r.e. non ricorsivo permettono la rappresentabilità di *ogni* insieme r.e., come avviene ad esempio per l'aritmetica di Peano.

Una domanda interessante diveniva così quella di sapere se esistono sistemi minimali nei quali tutte le funzioni ricorsive si possono rappresentare fortemente o debolmente. Una risposta significativa in questa direzione fu data nel 1950 da R. Robinson che isolò una teoria, la cosiddetta aritmetica $\mathfrak{Q}$ di Robinson, che è finitamente assiomatizzata e in cui tutte le relazioni ricorsive si possono rappresentare fortemente. Mette conto di riportare gli assiomi di questa teoria, il cui linguaggio contiene la costante $\bar{0}$, il simbolo $'$ per il successore, il segno di addizione $+$ e di moltiplicazione $\cdot$ e i cui assiomi sono i seguenti:

$Q1 \quad \forall x\, \forall y(x' = y' \rightarrow x = y),$
$Q2 \quad \forall x(\bar{0} \neq x'),$
$Q3 \quad \forall x(x \neq \bar{0} \rightarrow \exists y\, x = y'),$
$Q4 \quad \forall x(x + \bar{0} = x),$
$Q5 \quad \forall x\, \forall y(x + y') = (x + y)',$
$Q6 \quad \forall x(x \cdot \bar{0} = \bar{0}),$
$Q7 \quad \forall x\, \forall y(x \cdot y') = (x \cdot y) + x.$

L'assiomatizzabilità finita di Ω ci permette di ottenere immediatamente sfruttando il teorema di deduzione il teorema di Church come corollario della indecidibilità di Ω, ma l'importanza di Ω nasce anche da altre considerazioni, che riguardano il concetto stesso di insieme ricorsivo.

Già nel lavoro fondamentale del 1931, Gödel aveva sottolineato lo stretto rapporto tra rappresentabilità e ricorsività, enfatizzando nella scelta stessa del termine *entscheidungsdefinit* il carattere di dominabilità delle relazioni rappresentabili. Negli anni quaranta diversi ricercatori posero in luce come il collegamento non valesse in una direzione sola e come di fatto gli insiemi (come pure le funzioni e le relazioni) ricorsivi fossero caratterizzabili in termini di rappresentabilità. Fu Mostowski che in particolare insistette su questo collegamento, provando come tutte e sole le relazioni ricorsive sono rappresentabili nell'aritmetica di Peano e come tutti e soli gli insiemi r.e. siano debolmente rappresentabili nella stessa teoria. Il risultato fu successivamente esteso a tutte le estensioni finite di Ω. Si otteneva così una nuova caratterizzazione del concetto di ricorsività e di enumerabilità ricorsiva in termini di una particolare forma di definibilità all'interno di teorie formali.

La fecondità di questo approccio non tardò a emergere, quando Mostowski, in collaborazione con Grzegorczyk e Rill-Nardzewski provò che gli insiemi iperaritmetici si lasciavano caratterizzare in modo analogo come gli insiemi fortemente rappresentabili in un'estensione dell'aritmetica Ω di Robinson al secondo ordine ottenuta aggiungendo l'assioma di estensionalità

$$X = Y \longleftrightarrow \forall x(x \in X \longleftrightarrow x \in Y),$$

lo schema di comprensione

$$\forall X\, \exists x(x \in X \longleftrightarrow \mathscr{A}(x))$$

e l'assioma di induzione al secondo ordine.

Questi risultati sono particolarmente interessanti perché forniscono la base per generalizzazioni del concetto di ricorsivo e r.e. che emergeranno negli anni sessanta, una volta riformulati da G. Kreisel utilizzando il concetto di *definibilità invariante.* L'osservazione di Kreisel è che sostanzialmente non occorre far riferimen-

to a *tutti* i modelli della teoria Ω o delle sue estensioni, ma solo ad alcuni, osservando che in tutti modelli esiste una «copia» dei naturali: i denotati delle cifre. Si può allora dire che una relazione R sui naturali è *definita in modo invariante* relativamente ad una classe Γ di modelli della teoria Ω, quando esiste una formula $\mathscr{A}(x_1, ..., x_n)$, tale che per ogni naturale $a_1, ..., a_n$:

1) $\mathscr{A}(\bar{a}_1, ..., \bar{a}_n)$ o $\neg\mathscr{A}(\bar{a}_1, ..., \bar{a}_n)$ è vera in ogni modello di Γ e

2) $<a_1, ..., a_n> \in R$ se e solo se in ogni modello di Γ è vera $\mathscr{A}(\bar{a}_1, ..., \bar{a}_n)$.

Come Kreisel mostrava, si può verificare a questo punto che le relazioni ricorsive sono *esattamente* quelle definibili in modo invariante rispetto alla classe di tutti i modelli di Ω e, fatto più significativo, che le relazioni iperaritmetiche coincidono con quelle definibili in modo invariante nella classe di tutti gli ω-modelli che soddisfano l'estensione di Ω ottenuta con l'aggiunta di estensionalità e comprensione e quindi abbiano come dominio N.

Il ruolo dell'induzione viene spostato così sul concetto di ω-modello, cioè su quelle strutture del secondo ordine il cui insieme di individui è costituito dai numeri naturali e che soddisfano gli assiomi precedenti. Lavori successivi di Kreisel, di R. Gandy e W. Tait, negli anni sessanta, hanno approfondito l'analisi degli insiemi iperaritmetici come insiemi definibili in modo invariante, ad esempio introducendo sistemi come quello di sopra, in cui la comprensione è ristretta a formule che definiscono insiemi di questo tipo. Il fatto non stupisce se si considera che, come proposto da Kreisel nel 1960, è possibile identificare gli insiemi di naturali definibili predicativamente con gli iperaritmetici, o comunque anche senza accettare l'identificazione, come successivamente Kreisel tese a fare, esiste uno stretto legame fra i due concetti che la definizione di definibilità invariante pone in luce. Quello che ci preme sottolineare è che quanto detto sulla definibilità invariante ammette estensioni a teorie molto generali e permette quindi – come mostrato da Y. Moschovakis nel suo lavoro *Abstract Computability and invariant Definability* (*Computabilità astratta e definibilità invariante*, 1974) – di definire una nozione di ricorsività per strutture diverse dai naturali e offrire così un contesto in cui sviluppare in modo unitario varie generalizzazioni della ricorsività sorte at-

torno ai primi anni sessanta e su cui avremo occasione di ritornare.

Abbiamo parlato della caratterizzazione degli insiemi iperaritmetici e del loro collegamento con la nozione di definibilità predicativa. Per definizione, gli insiemi iperaritmetici si trovano al confine tra gli insiemi definibili utilizzando solo quantificazioni sui numeri naturali e che sono disposti nella gerarchia aritmetica, e quelli disposti nella gerarchia analitica, ottenibili quantificando anche su insiemi di tali numeri: in un senso quindi costituiscono il «limite» cui si approssimano gli insiemi definibili aritmeticamente. Negli anni cinquanta diversi ricercatori come M. Davis, A. Mostowski e S.C. Kleene mostrarono come anche gli iperaritmetici fossero disponibili in una gerarchia di progressiva «non costruttività» i cui livelli erano indiciati da un particolare segmento iniziale degli ordinali di seconda classe: gli ordinali costruttivi. Ciò permetteva l'estensione alla gerarchia iperaritmetica di problemi e nozioni impiegate studiando altre gerarchie e i lavori fondamentali su questo tema sono dovuti, oltre che a Kleene, a C. Spector. Non ci è possibile soffermarci su questi argomenti, ma vogliamo invece spendere qualche parola sugli ordinali costruttivi, il cui ruolo si dimostrò ben presto essenziale anche in altri tipi di indagine.

Come sappiamo l'idea di cercare una versione costruttiva del concetto di ordinale aveva trovato negli anni venti il suo protagonista soprattutto in Brouwer, ma gli ordinali costruttivi di cui parliamo furono introdotti nel 1936 da Church e Kleene sulla base della teoria della ricorsività come strumento per dare una notazione sistematica, in termini aritmetici, al segmento degli ordinali che si possono ottenere in maniera effettiva. Come sappiamo, gli ordinali si ottengono tutti da zero con due operazioni: passaggio al limite e successore. L'idea allora è di assegnare una *notazione* agli ordinali numerabili assegnando a ciascuno di essi un naturale che codifichi entro di sé le informazioni sufficienti per determinare mediante quali operazioni (successore o limite) è stato ottenuto da zero. La procedura è induttiva: al posto dell'ordinale 0 prenderemo il naturale 1 e se y è ordinale costruttivo, con 2^y indicheremo il suo successore e definiremo simultaneamente la relazione d'ordine $y < 2^y$. Nel caso degli ordinali limite, considereremo solo successioni definite da funzioni ricorsive, e porremo che se y è l'indice di una funzione ricorsiva che definisce una successione crescente, $3 \cdot 5^y$ sarà la notazione per l'ordinale limite della successione e si-

multaneamente stipuleremo che per ogni n sia $f(n) < 3 \cdot 5^y$, dove f è la funzione che definisce la successione. È chiaro che non tutti gli ordinali di seconda classe potranno esser costruttivi (avere cioè una notazione) in quanto gli ordinali di seconda classe sono più che numerabili, le notazioni numerabili. Esiste quindi un minimo ordinale non costruttivo che si indica con ω_1^c e si può provare che gli ordinali costruttivi coincidono esattamente col segmento determinato da ω_1^c e con quegli ordinali (detti anche *ricorsivi*) che hanno il tipo d'ordine di buoni ordinamenti ricorsivi dei naturali come provato da W. Markwald (1954) e Spector (1955).

Il ruolo centrale degli ordinali costruttivi, come mostra l'esempio della gerarchia iperaritmetica, è quello di costruire successioni costruttivamente dominabili; un esempio estremamente interessante di questo loro ruolo compariva nel 1939 nel già citato articolo di Turing sulle logiche ordinali. Qui Turing tentava una strada per superare le limitazioni poste in luce da Gödel che avrebbe avuto successivamente notevole fortuna e sarebbe stata estesa anche alla costruzione di sistemi predicativisti. L'idea di Turing era di definire progressioni di teorie, a partire dall'aritmetica, indiciando le successive estensioni in termini di ordinali costruttivi, il cui carattere di ordinale doveva essere provato nel caso della teoria $\mathfrak{T}_{\alpha+1}$ entro la teoria $\mathfrak{T}_\alpha$.

L'obiettivo era di superare l'ostacolo dell'incompletezza delle teorie considerando la dimostrabilità non entro *singole* teorie, ma lungo *successioni* di teorie. Turing dimostrava diversi risultati che collegavano il grado di completezza della successione con il tipo di ordinali necessari per indiciare la successione stessa. Come ricordato però, sarà negli anni sessanta che idee del genere si ripresenteranno tanto in connessione con i principi di riflessione e la costruzione di sistemi matematici predicativisti, quanto nell'analisi delle dimostrazioni di coerenza per sistemi più forti dell'aritmetica lungo le linee di Gentzen. In questo contesto furono studiati sistemi naturali per dare notazioni a segmenti di ordinali da K. Schütte e dalla sua scuola (riprendendo studi iniziati da Veblen nel 1908), da Solomon Feferman e da altri.

Per completare il panorama degli sviluppi che dopo gli anni trenta ha avuto la teoria della ricorsività, occorre accennare ad un tipo di ricerche che nel corso del tempo hanno acquistato un significato sempre maggiore non solo all'interno della teoria dell'effet-

tivo. Ci riferiamo alla teoria dei funzionali ricorsivi, le cui origini risalgono agli anni cinquanta. Un funzionale è un'operazione che si applica a funzioni e numeri per dare come valori numeri. Funzionali di ordine superiore ma di tipo finito sono quelli che possiamo definire induttivamente ponendo che un funzionale di tipo n si applica a funzionali di tipo $n-1$ e ha come valori numeri naturali. Esempi impliciti del concetto di funzionale si presentano spontaneamente nel contesto della teoria della ricorsività relativa: dire che la funzione f è ricorsiva nelle funzioni $f_1, ..., f_n$, significa sostanzialmente affermare che per ogni successione di argomenti $a_1, ..., a_k$ il suo valore si ottiene a partire da quelli delle f_i applicando gli schemi che generano le funzioni ricorsive. Se questa dipendenza dei valori di f da quelli delle f_i è uniforme negli argomenti numerici, essa ha un carattere schematico che ci permette di affermare che esiste un funzionale F tale che

$$F(f_1, ..., f_n, a_1, ..., a_k) = f(a_1, ..., a_k).$$

Funzionali di questo tipo vennero introdotti da Kleene nel 1952 e si chiamano *funzionali ricorsivi parziali*. Per essi è possibile sviluppare una teoria – come Kleene mostrò – che è strettamente analoga a quella delle funzioni ricorsive, provando ad esempio un teorema di forma normale. Questo esempio di estensione del concetto di computabilità ai funzionali presenta alcuni aspetti fondamentali che è possibile generalizzare a classi più ampie di funzionali. Come Kleene mostrò, se $F(\alpha, x)$ è un funzionale ricorsivo parziale che prende valore y sugli argomenti dati, si può verificare il fatto che esso sarà *compatto* nel senso che esisterà una funzione finita (cioè con dominio finito) $u \subseteq \alpha$, per cui $F(u, x) = y$; esso inoltre sarà *monotono* nel senso che se β è un'estensione della funzione parziale α allora avremo sempre che $F(\beta, x) = y$. Le due proprietà intuitivamente significano che computando F per gli argomenti α ed x noi useremo solo una *parte finita* di valori che α può assumere e quindi una sottofunzione ristretta ad una parte finita del suo dominio: in altre parole ancora, ogni computazione sulla base dello «schema» F costituisce un albero finito in cui si usa una quantità finita di informazione. La monotonia è un'ulteriore controparte di questo fatto, e la congiunzione delle due proprietà, compattezza e monotonia, si può esprimere dicendo che il funzionale è *continuo* in

quanto, data un'opportuna topologia allo spazio delle funzioni, la controimmagine di un aperto rispetto ad F è ancora un aperto.

Ritroviamo così quel legame tra continuità e costruttività che già Brouwer aveva considerato. Il fatto importante è che la continuità ha come corollario l'esistenza di *punti fissi* per i funzionali, funzioni α per cui $F(\alpha, x) = \alpha(x)$. Questo significa che possiamo definire una funzione (si ricordi, parliamo sempre di funzioni parziali) scrivendo condizioni che i suoi valori devono soddisfare e in cui essa stessa può comparire. È questo che afferma il famoso primo teorema di ricorsione dimostrato da Kleene nel 1952 e che dice che ogni funzionale parziale ricorsivo F ammette un minimo punto fisso che è una funzione ricorsiva, una funzione parziale α, tale che

$$\forall x(\alpha(x) = F(\alpha, x))$$

$$\forall x(\beta(x) = F(\beta, x)) \rightarrow \alpha \subseteq \beta.$$

Questo mostra che la continuità ci garantisce la validità di uno schema di definizione per funzioni computabili estremamente potente. È su queste basi che nel 1959, per vie indipendenti, M. Davis, Kleene e Kreisel introdussero il concetto di *funzionale computabile di tipo finito*, basandosi sull'idea che è proprio la continuità il veicolo centrale della computabilità. Kreisel utilizzò i funzionali continui nello studio dell'aritmetica intuizionista portando avanti quelle ricerche sull'isolamento del significato costruttivo dei teoremi dell'aritmetica e dell'Analisi che l'avevano condotto già nel 1951 a definire la cosiddetta *no counterexemple interpretation*; questa estendeva l'idea delle forme normali di Herbrand e cercava di esprimere, in termini di esistenza di funzionali di tipo finito, la non esistenza di controesempi che falsificano una data formula. Questa non esistenza però non andava intesa in senso astratto, ma costruttivo cosicché – come Kreisel dimostrò in diversi casi – la dimostrabilità di un enunciato all'interno di date teorie si poteva caratterizzare in termini delle proprietà dei funzionali che «neutralizzano» i controesempi, e più «elementari» sono le teorie più «costruttivi» sono i funzionali.

Altro campo cui la teoria dei funzionali continui venne presto applicata è l'Analisi ricorsiva, sviluppata lungo linee differenti da R. Goodstein, E. Specker, A. Grzegorczyk, S. Mazur ed altri in

cui i reali ricorsivi sono identificati con funzioni ricorsive sui numeri naturali e le funzioni reali con funzionali. Ma questa non è l'unica via tentata per introdurre il concetto di funzionale computabile. Nel caso dei funzionali di cui sinora abbiamo parlato, le computazioni sono alberi finiti, ma come mostrò Kleene in due lavori fondamentali rispettivamente del '59 e del '63, è possibile sviluppare una teoria della computabilità per funzionali di tipo finito facendo cadere la compattezza e ammettendo alberi di computazione che hanno rami finiti ma ammettono diramazioni infinite. In questa prospettiva le computazioni non possono essere più codificate da indici naturali e cadono quindi teoremi come quello di forma normale. L'idea di Kleene è allora di definire la classe dei funzionali ricorsivi mediante schemi che simultaneamente forniscano gli indici necessari. Col tempo sono emersi collegamenti fra i due tipi di generalizzazione ma quello che a noi interessa è che la strada seguita da Kleene ammette generalizzazioni più astratte che verranno sviluppate negli anni sessanta, sempre fondandosi sull'idea di introdurre schemi che postulano l'esistenza di indici senza la necessità di doverli definire.

Vedremo più avanti come nel 1958 un'altra generalizzazione del concetto di computabilità per funzionali di tipo finito sia stata presentata da Gödel con i suoi funzionali ricorsivi primitivi. Questi funzionali svolgeranno un ruolo fondamentale nello studio della matematica classica e intuizionista. Negli anni sessanta tutte queste diverse analisi di come estendere a oggetti diversi dalle funzioni, quali i funzionali, il concetto di computabilità basandosi sulla decidibilità, la finitezza, la struttura delle computazioni, la definizione di macchine ecc., portarono all'esigenza di un approccio più astratto che da una parte permettesse una visione unitaria delle diverse generalizzazioni, dall'altra ponesse meglio in rilievo gli aspetti salienti della stessa teoria delle funzioni ricorsive (ORT).

A. Mostowski, *Thirty Years of foundational Studies*, Blackwell, Oxford 1966.

J. Herbrand, *Ecrits logiques*, a cura di J. van Heijenoort, PUF, Parigi 1968.

J. van Heijenoort, *Selected essays*, Bibliopolis, Napoli 1985.

K. Gödel, *Collected Works*, a cura di vari autori, Oxford University Press, New York. Sono apparsi finora 2 volumi (1986, 1990) che comprendono tutte le opere edite.

AA.VV., *Il teorema di Gödel. Una messa a fuoco*, a cura di S.G. Shanker, Muzzio, Padova 1991.

A.S. Troelstra e D. van Dalen, *Constructivism in Mathematics*, 2 volumi, North Holland, Amsterdam 1988.

AA.VV., *The L.E.J. Brouwer Centenary Symposium*, a cura di A.S. Troelstra e D. van Dalen, North Holland, Amsterdam 1982.

A.S. Troelstra, *Principles of Intuitionism*, Springer Verlag, Berlino 1969.

G. Gentzen, *The collected papers*, a cura di M.E. Szabo, North Holland, Amsterdam 1969.

AA.VV., *Teoria della dimostrazione*, a cura di Donatella Cagnoni, Feltrinelli, Milano 1981.

E. Moriconi, *La teoria della dimostrazione di Hilbert*, Bibliopolis, Napoli 1987.

A.C. Leisenring, *Mathematical Logic and Hilbert symbol*, MacDonald, Londra 1969.

S.C. Kleene, *Introduction to Metamathematics*, North Holland, Amsterdam 1952.

AA.VV., *The Undecidable*, a cura di M. Davis, Raven Press, New York 1965.

A.A. Markov, *Theory of Algorithms*, Israel Program for Scientific Translations, Gerusalemme 1962 [edizione originale 1954].

H. Hermes, *Enumerabilità, decidibilità, computabilità*, Boringhieri, Torino 1975 [edizione originale 1961].

A. Church, *The calculi of lambda conversion*, Princeton University Press, Princeton 1941.

H. Rogers, *Teoria delle funzioni ricorsive e della computabilità effettiva*, Tecniche nuove, Milano 1992 [edizione originale 1967].

P. Odifreddi, *Classical recursion Theory*, North Holland, Amsterdam 1989.

A. Hodges, *Storia di un enigma. Vita di Alan Turing*, Bollati Boringhieri, Torino 1991 [edizione originale 1983].

AA.VV., *Recursive Function Theory*, a cura di J.C.E. Dekker, AMS, Providence 1962.

R. Smullyan, *Theory of formal Systems*, Princeton University Press, Princeton 1961.

CAPITOLO SESTO
DOPO LA SECONDA GUERRA MONDIALE

1. INTRODUZIONE

Nel nostro schema generale avevamo riconosciuto che la ricerca logica, dopo che nei primi trent'anni del secolo si era mossa sostanzialmente articolandosi in *programmi* (scuole), aveva subito proprio negli anni trenta – e in virtù di tutta una serie di risultati che abbiamo sopra descritto – una radicale trasformazione, indirizzandosi verso l'approfondimento di specifici *temi*, pur se di portata generale, che erano appunto scaturiti dalle stesse indagini programmatiche. Avevamo anche avvertito però che questo non significava impegnarsi in una «pratica senza principi», ma rappresentava piuttosto il tentativo di saggiare sul terreno concreto delle applicazioni e delle analisi particolari la portata degli strumenti e degli approcci precedentemente elaborati. Proprio per questo ci era sembrato opportuno, e anzi necessario, soffermarci ad analizzare abbastanza dettagliatamente il passaggio dalle concezioni «preformali» a quelle «formali», tentando di metterne in luce le motivazioni generali da una parte e gli specifici strumenti di concettualizzazione dall'altra; convinti come siamo che la comprensione di tali concetti cardine sia *essenziale* per poter penetrare quella che è l'attuale situazione della ricerca, anche nella sua espressione più propriamente «tecnica».Volgendoci ora agli sviluppi dell'indagine logica dopo il secondo conflitto mondiale, non sarà più possibile attenersi ad una procedura analoga a quella sino ad ora seguita e articoleremo il discorso più che per scuole per filoni di ricerca. D'altra parte il panorama è così ramificato e coinvolge contatti così stretti con la pratica matematica, che per descriverlo puntualmente sarebbe necessario un lavoro preliminare che non ci sembra plausibile richiedere né alle nostre forze né a quelle del lettore. Ci limiteremo quindi ad illustrare le vie di svi-

luppo più generali, enfatizzando quelli che secondo noi sono gli elementi di novità. Sul piano temporale ci porremo *grosso modo* come limite i primi anni settanta, e questo per ragioni che abbiamo già chiarito nell'introduzione. Visto il sempre più stretto intrecciarsi dei diversi rami della logica matematica, teoria degli insiemi, dei modelli, della dimostrazione, ecc., saremo spesso costretti a procedere in modo non uniforme: in alcuni casi tratteremo direttamente in modo unitario le indagini che vanno dalla fine della guerra fino agli anni settanta (come d'altra parte abbiamo già fatto nel capitolo precedente per la teoria della ricorsività e nel secondo capitolo per la logica algebrica) in altri procederemo in due tappe, usando come discrimine sostanzialmente gli anni sessanta. Prenderemo le mosse dalla teoria dei modelli, che della ricerca logica postbellica rappresenta alcune delle tendenze più caratteristiche.

2. LO SVILUPPO DELLA TEORIA DEI MODELLI

Come già si è detto nel paragrafo 3 del capitolo precedente, la definizione tarskiana di verità, ponendo le basi della semantica dei linguaggi formalizzati ed introducendo l'esigenza di metodi infinitari nella metamatematica, apriva la strada ad un nuovo modo di vedere lo studio delle teorie, un modo che instaurava legami più diretti tra la concreta pratica matematica e l'analisi logica. Era in questo spirito che la metodologia delle scienze deduttive propugnata da Tarski attorno agli anni trenta si era proposta lo studio delle teorie formalizzate dando così inizio all'esame delle loro proprietà, sia sintattiche che semantiche, senza porsi preclusioni programmatiche sui metodi dimostrativi. Come abbiamo visto, con questo si realizzava una rottura esplicita con l'idea hilbertiana di metamatematica concepita come studio, in base a metodi rigidamente finitisti, delle pure proprietà sintattiche, per dare l'avvio ad un'analisi «matematica» delle teorie, analisi che trovava la sua collocazione più naturale non più entro una matematica ristretta quale quella finitista, ma all'interno della più generale teoria degli insiemi. In tal modo lo studio delle teorie formalizzate veniva a costituire, come l'algebra o la topologia generale, una delle articolazioni di quella matematica astratta e «infinitaria» che proprio

negli anni trenta, attraverso l'impiego sistematico dei più potenti strumenti insiemistici (assioma di scelta, ipotesi di misurabilità, del continuo ecc.), andava dando ad algebra e topologia una nuova forma, conferendo loro una generalità sino ad allora mai vista.

Se sulle prime, si ricorderà, il rapporto fu sostanzialmente unidirezionale, risolvendosi nel tentativo (largamente riuscito) di applicare metodi algebrici e topologici allo studio delle teorie formalizzate, ben presto esso si capovolse, aprendo prospettive inaspettate. La semantica tarskiana offriva infatti la possibilità di compiere il passaggio inverso, di applicare cioè concetti e metodi metamatematici alla risoluzione di problemi algebrici e topologici. È dal tentativo di sviluppare appieno questa possibilità e di analizzarne le modalità che ha avuto origine la moderna teoria dei modelli nel suo duplice aspetto di studio delle teorie formalizzate in rapporto ai relativi modelli e di applicazione di concetti metamatematici a specifici problemi algebrici o in generale matematici.

Alla base di questo rapporto stava un fatto che abbiamo già avuto occasione di sottolineare. Le realizzazioni, i possibili modelli delle teorie formalizzate che sono al centro dell'indagine logica, sono riconducibili a strutture, algebre o sistemi relazionali, dello stesso genere dei gruppi, anelli, campi, ecc., oggetti di studio dell'algebra e più in generale della matematica e l'algebra universale di Birkhoff e Ore aveva ampiamente sottolineato l'unità che sottostava a tutte queste ricerche. In tutti i casi si tratta di insiemi su cui sono definite relazioni e funzioni le cui proprietà definitorie sono fissate per via assiomatica. La differenza fra approccio algebrico e logico stava nell'enfasi che veniva posta sui due aspetti che concorrono alla definizione delle strutture; l'algebrista ne sottolineava l'aspetto insiemistico e si concentrava quindi sull'indagine delle possibili operazioni che su di esse si possono definire e sui morfismi che tra di esse sussistono; il logico invece privilegiava l'altro aspetto, il carattere assiomatico della caratterizzazione delle relazioni e funzioni che definiscono le strutture e si concentrava quindi sullo studio delle teorie formalizzate.

L'obiettivo di fondo della teoria dei modelli è quello di mediare questi due momenti, più precisamente quello di analizzare i rapporti tra la *forma* linguistica dei sistemi d'assiomi e le *proprietà insiemistiche* dei modelli (ed è in questo senso che ci si riallaccia e si estende la semantica come intesa precedentemente). In questo

modo si raggiungono due scopi: da una parte si può sfruttare tutto quanto si sa sulle strutture e le operazioni su di esse definite nello studio delle teorie formalizzate, dall'altra risulta possibile affrontare problemi algebrici ricorrendo allo studio delle teorie formalizzate che determinano le varie classi di strutture, in generale, alle proprietà dei linguaggi. I due aspetti si completano a vicenda, scambiandosi mutuamente strumenti e problemi. L'algebra offre al logico metodi di costruzione di strutture per affrontare lo studio delle teorie e simultaneamente stimola la elaborazione di strumenti linguistici sempre più adeguati ai problemi algebrici; la logica offre all'algebrista informazioni sulle proprietà dei sistemi assiomatici e stimola la definizione di metodi per la costruzione di strutture che meglio riflettano i fatti riguardanti la loro caratterizzazione linguistica. La stessa ragion d'essere della teoria dei modelli porta a un intrecciarsi continuo delle due impostazioni, così che è difficile porre una netta linea di demarcazione anche solo a livello di metodi; procedimenti costruttivi e non si innestano gli uni negli altri senza soluzione di continuità e le restrizioni finitiste (retaggio della concezione hilbertiana della metamatematica) lasciano sempre più posto all'impiego sistematico di principi insiemistici altamente non costruttivi quali l'assioma di scelta, l'ipotesi del continuo, ecc.

Si realizza così in pieno quell'ideale della metamatematica come studio *globale* delle proprietà delle teorie formalizzate di cui la semantica e la metodologia delle scienze deduttive tarskiane erano state le avvisaglie. È in questo senso che all'inizio degli anni cinquanta, quasi simultaneamente Tarski in una serie di articoli dal titolo appunto *Contributions to the Theory of Models* (*Contributi alla teoria dei modelli*, 1954) che riprendono una comunicazione del 1950 *Some Notions and Methods on the Borderline of Algebra and Metamathematics* (*Nozioni e metodi al confine tra algebra e metamatematica*), L. Henkin nel 1953 in *Some Interconnections between modern Algebra and Mathematical Logic* (*Interconnessioni fra algebra moderna e logica matematica*) e soprattutto Abraham Robinson (1918-74) con la sua tesi pubblicata nel 1951 dal titolo *On the Metamathematics of Algebra* (*Sulla metamatematica dell'algebra*), presentavano un vero e proprio programma di studio sistematico dei modelli delle teorie formali al di fuori di un orizzonte strettamente fondazionale; l'enfasi era diversa: Tarski aveva più l'occhio all'algebra universale nel senso di Birkhoff,

mentre Robinson, e come lui Henkin, vedeva nella teoria dei modelli un modo per analizzare in generale procedure e concetti tipici della matematica e in particolare dell'algebra e della geometria.

Non è un caso, come testimoniano questi lavori, che il primo affermarsi della teoria dei modelli coincida con l'estensione del concetto stesso di linguaggio formalizzato e con la parallela generalizzazione dei teoremi semantici fondamentali – di compattezza e di Löwenheim-Skolem – nel tentativo di farne strumenti adeguati alle applicazioni algebriche. Nella loro naturalezza queste estensioni avevano un carattere decisamente rivoluzionario in quanto segnavano un abbandono senza possibilità di recupero delle restrizioni finitiste di tradizione hilbertiana ed aprivano la strada all'applicazione di concetti e metodi esplicitamente infinitari.[1] È in due articoli del logico sovietico Anatolij Malcev, rispettivamente del 1936 e del 1941, che questo passaggio – come si ricorderà – era stato effettuato nella sua forma più netta. Con questi lavori Malcev forniva simultaneamente gli strumenti centrali ed i primi esempi significativi di un'applicazione della metamatematica all'algebra, strumenti ed esempi che per lungo tempo sarebbero rimasti paradigmatici ed ispiratori, una volta che verso gli anni cinquanta furono conosciuti anche in Occidente.

La prima innovazione introdotta, come già si è detto, riguarda la nozione stessa di linguaggio formalizzato. Sino ad allora, in conformità al punto di vista hilbertiano, un linguaggio $\mathscr{L}$ del primo ordine era stato determinato oltre che dalle ordinarie costanti logiche, connettivi e quantificatori, da una famiglia *numerabile* di costanti extralogiche, classificate in costanti individuali, predicative e funzionali. Questo affinché la collezione delle formule e degli enunciati risultasse decidibile e quindi suscettibile di studio entro la matematica finitaria. Questa condizione vincolante risulta però non solo arbitraria da un punto di vista astratto, ma addirittura inaccettabile una volta che si voglia, data una qualsiasi struttura $\mathfrak{M}$, descriverla «il più adeguatamente possibile» a livello lin-

[1] Naturalmente, c'è da tener presente l'altro tipo di superamento di tali limiti che abbiamo visto muoversi sulla linea di Gentzen in un contesto che si può dire di mediazione fra sintassi e semantica, anche se è stato tuttavia sviluppato in termini della teoria della dimostrazione. Va osservato comunque che questo tentativo rimane pur sempre nell'ambito del costruttivo.

guistico. Per far questo occorre possedere nomi per ogni individuo di $\mathfrak{M}$ così da poter dire per ognuno di essi quali proprietà (espresse nel linguaggio) siano vere o no. Le costanti individuali di un linguaggio, però, possono essere al più numerabili e quindi una simile «descrizione» risulta impossibile per strutture arbitrarie, la cui definizione non implica affatto restrizioni di cardinalità.

È sulla base di queste considerazioni che Malcev introduce linguaggi del primo ordine con un numero arbitrario di costanti extralogiche e ne analizza le proprietà semantiche fondamentali, giungendo così a stabilire quel teorema di compattezza o *principio di localizzazione* che costituisce ancor oggi uno degli strumenti fondamentali della teoria dei modelli. Il principio, per la cui dimostrazione è essenziale l'uso dell'assioma di scelta, era una generalizzazione del teorema di compattezza di Gödel e afferma che, dato un insieme K di enunciati del primo ordine, avente cardinalità *arbitraria*, esso ha un modello se e solo se ogni suo sottoinsieme finito ha un modello. In questo modo è possibile ricondurre l'esistenza di modelli per insiemi infiniti arbitrari all'esistenza di modelli per insiemi finiti. Poggiandosi su questo principio, Malcev riusciva a dimostrare in modo estremamente elegante ed uniforme un numero notevole di teoremi sui gruppi infiniti e mettendone in luce l'aspetto comune e la dipendenza dal principio di localizzazione faceva vedere, per usare le sue stesse parole, come «essi non fossero specificatamente algebrici ma potessero essere ottenuti come conseguenze immediate di enunciati generali della logica matematica». In termini netti e precisi, veniva così formulato e realizzato un obiettivo centrale della moderna teoria dei modelli.

Il metodo seguito da Malcev è estremamente diretto e costituisce un paradigma che ha trovato un numero incredibile di applicazioni; l'idea di fondo è quella di ricondurre l'esistenza di strutture aventi date proprietà algebriche alla esistenza di modelli per dati insiemi di enunciati. Il problema diviene quindi quello di trovare per le proprietà algebriche in oggetto un insieme di enunciati del primo ordine che le «riflettano», in altri termini tale che una struttura sia modello dell'insieme se e solo se gode delle proprietà in questione; diremo *elementare* una proprietà siffatta. Una volta fatto questo, il principio di localizzazione da criterio per l'esistenza di modelli si trasforma in criterio per l'esistenza di strutture con le proprietà volute. Visti in quest'ottica, numerosi teoremi al-

gebrici si rivelano immediate conseguenze del principio di localizzazione. È questo che Malcev mostrava nel caso dei teoremi locali sui gruppi (che noi abbiamo ricordato nel paragrafo v, 2), teoremi che affermano che un gruppo ha una data proprietà P se e solo se ogni suo sottogruppo finitamente generato ne gode. Se la proprietà P è elementare, la cosa diviene infatti immediata seguendo il metodo di Malcev.

Dato un gruppo $\mathfrak{G}$, indichiamo con $D(\mathfrak{G})$ quello che oggi, seguendo una terminologia introdotta nel 1950 da Abraham Robinson, che introdusse tale concetto, viene detto *diagramma* di $\mathfrak{G}$ ossia l'insieme di tutte le formule atomiche o negazioni di atomiche vere in $\mathfrak{G}$ e contenenti, oltre a un simbolo che denota l'operazione del gruppo e al simbolo d'identità, nomi per ogni elemento di $\mathfrak{G}$. Consideriamo allora l'insieme K di enunciati che, per ipotesi, «riflette» la proprietà P. Sia quindi $K' = K \cup D(\mathfrak{G}) \cup K_{\mathfrak{G}}$, dove con $K_{\mathfrak{G}}$ indichiamo l'insieme di assiomi per la teoria dei gruppi, che sappiamo essere formulabile al primo ordine. È chiaro che ogni modello di K', in quanto modello di $D(\mathfrak{G})$ sarà estensione (eventualmente impropria) di $\mathfrak{G}$ ed in quanto modello di K godrà di P; d'altra parte ogni sottogruppo di $\mathfrak{G}$, avrà come diagramma un sottoinsieme di $D(\mathfrak{G})$. In base al principio di localizzazione non esisterà modello di K' solo nel caso che un sottoinsieme finito di K' ne sia privo; ma allora è facile concludere che se ogni sottogruppo finitamente generato di $\mathfrak{G}$ gode di P, cioè è modello di K, esisterà un modello di K', quindi esisterà una estensione di $\mathfrak{G}$ che gode di P. Basta ora che la proprietà in questione sia «ereditaria» (tale cioè che se goduta da una struttura sia goduta anche dalle sue sottostrutture) per concludere che $\mathfrak{G}$ stesso godrà di P. In questo modo, ad esempio, possiamo dimostrare che un gruppo è ordinabile se e solo se lo è ogni suo sottogruppo finitamente generabile; che contiene un sottogruppo commutativo di indice $\leqslant n$ se e solo se questo vale per ogni suo sottogruppo finitamente generato e così via. Gli esempi si possono moltiplicare e, come è facile vedere, non sono affatto limitati ai gruppi; risultati analoghi si ottengono nel caso di anelli, campi, algebre di Boole, ecc.; l'essenziale è che tanto le proprietà algebriche quanto il tipo di strutture che si analizzano siano descrivibili entro il primo ordine utilizzando all'occorrenza costanti extralogiche non occorrenti nella definizione delle strutture in esame.

Malgrado le ovvie possibilità che aprivano, le opere di Malcev rimasero per lungo tempo senza seguito, sino agli anni attorno al 1950 quando, per vie indipendenti, Tarski, A. Robinson e L. Henkin pubblicarono i lavori di cui si è detto nei quali venivano introdotti quelli che sono i veri e propri concetti fondamentali della attuale sistemazione della teoria dei modelli. Alla base di questa elaborazione rimaneva come strumento dimostrativo fondamentale il principio di localizzazione, riscoperto da Robinson e Henkin. I nuovi concetti introdotti permettevano però un'articolazione molto più duttile; accanto al concetto di diagramma, introdotto da Robinson, e di cui abbiamo parlato, l'altro è quello di *estensione elementare* introdotto da Tarski e Robert Vaught nel 1956. Data una struttura $\mathfrak{A}$, una sua sottostruttura $\mathfrak{B}$ si dice sottostruttura *elementare* (e $\mathfrak{A}$ estensione *elementare* di $\mathfrak{B}$), e si scrive $\mathfrak{B} \preccurlyeq \mathfrak{A}$ quando ogni enunciato – formulato in un linguaggio che oltre a nomi per tutte le funzioni e relazioni di $\mathfrak{A}$ contiene anche nomi per tutti gli individui di $\mathfrak{B}$ – è vero in $\mathfrak{A}$ se e solo se è vero in $\mathfrak{B}$, vale a dire tutti gli elementi di $\mathfrak{B}$ posseggono relativamente a $\mathfrak{B}$ le stesse proprietà elementari che posseggono rispetto ad $\mathfrak{A}$.

Il concetto si rivelò di importanza centrale. Tarski e Vaught riuscirono a dimostrare un utile criterio per l'esistenza di questa relazione fra strutture: date due strutture $\mathfrak{A}$ e $\mathfrak{B}$ di cui la seconda sia sottostruttura della prima, $\mathfrak{A}$ è estensione elementare di $\mathfrak{B}$ se e solo se dato un qualsiasi enunciato esistenziale $\mathscr{A} = \exists y \mathscr{B}(y)$, formulato nel linguaggio di $\mathfrak{B}$ e vero in $\mathfrak{A}$, esiste un elemento a di $\mathfrak{B}$ tale che $\mathscr{B}(a)$ è vero in $\mathfrak{A}$. L'importanza del concetto sta nel fatto che esso permette di articolare un approccio puramente semantico alle classificazioni delle teorie. Infatti, caso particolare della nozione di estensione elementare è quello di *equivalenza elementare*: due strutture $\mathfrak{A}$ e $\mathfrak{B}$ si dicono elementarmente equivalenti (e si scrive $\mathfrak{A} \equiv \mathfrak{B}$) quando sono indiscernibili dal punto di vista linguistico, cioè ogni enunciato del linguaggio di base vero in $\mathfrak{A}$ è necessariamente vero in $\mathfrak{B}$ e viceversa. È evidente che $\mathfrak{A}$ e $\mathfrak{B}$ sono elementarmente equivalenti quando una delle due è estensione elementare dell'altra; non vale invece evidentemente il viceversa. Si vede subito che il concetto di teoria completa ammette una formulazione esclusivamente in termini di modelli: una teoria $\mathfrak{T}$ è completa se e solo se tutti i suoi modelli sono elementarmente equivalenti. Questo criterio risultava molto importante in quanto per-

metteva di determinare la completezza di una teoria esclusivamente in base allo studio dei suoi modelli e risultava molto più duttile di un analogo criterio dimostrato da Vaught nel 1954 e secondo il quale se una teoria ammette solo modelli infiniti e per un certo cardinale $\mathfrak{m}$ tutti i modelli di cardinalità $\mathfrak{m}$ sono fra loro isomorfi (si dice in questo caso che la teoria è *categorica in potenza* $\mathfrak{m}$) allora la teoria è completa. La maggior duttilità permessa dal concetto di equivalenza elementare nasce dal fatto che quest'ultimo è meno restrittivo di quello di isomorfia, nel senso che due strutture $\mathfrak{A}$ e $\mathfrak{B}$ isomorfe sono anche elementarmente equivalenti ma non viceversa, come insegna il teorema di Skolem.

Nella pratica, è spesso attraverso lo studio di estensioni o strutture elementari che si stabilisce quando due strutture sono elementarmente equivalenti. Dire che $\mathfrak{A} \preccurlyeq \mathfrak{B}$ significa infatti che gli individui di $\mathfrak{A}$ visti in $\mathfrak{A}$ o in $\mathfrak{B}$ hanno le stesse proprietà elementari. Il fatto che per ogni coppia di modelli $\mathfrak{A}$, $\mathfrak{B}$ per cui $\mathfrak{A} \preccurlyeq \mathfrak{B}$ si abbia anche $\mathfrak{B} \equiv \mathfrak{A}$ ha un grande significato per una teoria, in quanto vuol dire che gli elementi di un modello hanno proprietà elementari che sono indipendenti dallo specifico ambiente in cui li situiamo (purché, ovviamente, si tratti sempre di modelli della teoria). Collegandosi ad alcune situazioni algebriche centrali – come il teorema degli zeri di Hilbert – nel 1949 A. Robinson introduceva il concetto di *model-completezza* (model-completeness) che analizza appunto questi tipi di teorie.

Una teoria $\mathfrak{T}$ è model-completa quando, dato un suo qualunque modello $\mathfrak{A}$, ogni altro modello $\mathfrak{B}$ di $\mathfrak{T}$ che sia estensione di $\mathfrak{A}$ è *estensione elementare* di $\mathfrak{A}$. Robinson forniva criteri efficaci per la model-completezza e mostrava che ogni teoria $\mathfrak{T}$ model-completa che possegga un modello primo (ossia un modello $\mathfrak{M}$ tale che ogni altro modello di $\mathfrak{T}$ possegga una sottostruttura isomorfa con $\mathfrak{M}$) è completa. Con questo criterio Robinson riusciva a dare una prima interpretazione in termini puramente modellistici del cosiddetto metodo dell'eliminazione dei quantificatori introdotto da Skolem e usato successivamente da H. Langford e soprattutto da Tarski, ottenendo fra gli altri come corollario una nuova dimostrazione di completezza per la teoria dei campi algebricamente chiusi di caratteristica data e per i campi reali chiusi che Tarski aveva appunto raggiunto con quel metodo.

Tutte queste tecniche per stabilire la completezza di teorie (mo-

del-completezza, eliminazione dei quantificatori, ecc.) hanno immediati corollari per quanto riguarda la *decidibilità*, come già si è detto parlando del teorema di Janiczack. Se una teoria assiomatizzabile è completa è infatti anche decidibile e questo spiega il particolare interesse dei metodi in questione. Già prima dei risultati di Tarski, M. Presburger, ad esempio, aveva provato la decidibilità del frammento dell'aritmetica di Peano che riguarda la sola addizione e negli anni venti – come ricordato – Langford aveva studiato sistematicamente, attraverso l'eliminazione dei quantificatori, varie teorie dell'ordine lineare mostrando la completezza della teoria degli ordini densi senza estremi o, con estremi, forme deboli di completezza per teorie più generali, ecc. Obiettivo del metodo dell'*eliminazione dei quantificatori* (E.Q.) è di vedere se all'interno di una data teoria $\mathfrak{T}$ ogni formula $\mathscr{A}(x_1,...,x_n)$ è dimostrabilmente equivalente ad un'altra formula $\mathscr{A}^*(x_1,...,x_n)$ *priva di quantificatori*, nel senso che si ha che

$$\mathfrak{T} \vdash \forall x_1,...,x_n(\mathscr{A}(x_1,...,x_n) \longleftrightarrow \mathscr{A}^*(x_1,...,x_n)).$$

L'eliminabilità dei quantificatori è una proprietà estremamente interessante che si può provare sia per via sintattica – come fatto da Tarski nei lavori citati – sia con metodi semantici, come mostrava in modo sistematico A. Robinson.

Nel caso di molte teorie algebriche infatti la E.Q. è collegata a criteri riguardanti la risolubilità dei sistemi di equazioni (come, nel caso dei campi reali chiusi, il teorema di Sturm) ed ha conseguenze algebriche assai interessanti, anche indipendentemente da problemi di decibilità; queste emersero solo verso gli inizi degli anni sessanta, quando si indagarono più da vicino con metodi semantici i suoi legami con il più debole concetto di model-completezza utilizzando anche i modelli saturi.

I lavori di Tarski sui campi reali chiusi – che risalgono agli anni trenta ma furono resi noti più tardi – avevano invece conseguenze interessanti sul piano fondazionale per quel che riguarda la geometria elementare. Tarski studiava infatti sistemi di geometria piana, proiettiva, affine, euclidea, prendendo – nell'ultimo caso – come primitivi i concetti di *stare tra* e di *equidistanza* tra coppie di punti $E\ (x, y, z, v)$ (che non è altro che la *congruenza* fra i segmenti che corrispondono alle coppie di punti x, y e z, v).

Dopo aver mostrato che ogni modello della teoria per il piano euclideo è isomorfo al piano costruito su un campo ordinato reale chiuso, Tarski poteva *interpretare* (come nella geometria analitica ordinaria) gli enunciati riguardanti la geometria in termini di enunciati sui campi reali chiusi e concludere che la geometria euclidea *elementare* (formulata cioè nel linguaggio elementare di cui si è detto) del piano è decidibile e completa. Lo stesso si poteva fare – applicando tecniche analoghe – per quella affine e per quella proiettiva. Negli anni cinquanta i risultati in questa direzione si moltiplicarono soprattutto per opera di Tarski e della sua scuola (Wanda Szmielew in particolare) mostrando come le tecniche della teoria dei modelli costituissero uno strumento centrale nello studio della metamatematica di teorie classiche come la geometria.

L'idea di *interpretare* una teoria in un'altra (come la geometria nella teoria dei campi reali chiusi) in modo che ogni enunciato $\mathscr{A}$ fosse dimostrabile nella prima se e solo se la sua «interpretazione» $\mathscr{A}^*$ era teorema della seconda, permetteva d'altra parte di ottenere anche molti risultati significativi di indecidibilità, interpretando l'aritmetica di Peano o una sua sottoteoria come Ω in altre teorie (nel caso più semplice, relativizzando le formule) che risultavano così indecidibili: questo era il caso per la teoria dei gruppi (Tarski, 1946) per quella degli anelli e dei domini d'integrità (R. Robinson, 1951), ecc. Le tecniche utilizzate, come molti dei risultati, costituiscono il nucleo di un libro, *Undecidable Theories* (*Teorie indecidibili*), pubblicato nel 1953 da Tarski, R. Robinson e A. Mostowski, un classico sull'argomento.

Eliminazione dei quantificatori, completezza e decidibilità costituirono a partire da questi anni uno dei campi in cui i metodi logici avevano un impatto più diretto con la pratica matematica. Questo vale per la geometria algebrica (già nel 1949 Tarski sottolineava come il teorema di Chevalley sulle proiezioni di varietà affini fosse una conseguenza della E.Q. per i campi algebricamente chiusi), la geometria algebrica *reale* (la geometria degli spazi R^n dove i punti hanno coordinate reali) in cui – come è stato mostrato in tempi recenti da studiosi come M. Coste, G. Brumfel e altri – la logica ha un grosso ruolo da svolgere e infine la teoria dei moduli, delle algebre, delle estensioni dei campi, che costituiscono uno degli aspetti più significativi della ricerca logica orientata verso la

pratica matematica che si è andata costituendo sull'esempio dei lavori di Tarski e soprattutto di A. Robinson a partire dagli anni cinquanta. Di alcuni di questi sviluppi avremo occasione di parlare più avanti, quando altri elementi avranno fatto la loro comparsa nel nostro quadro.

Volgiamoci ora ad un altro tipo di problemi, particolarmente vivo in questi anni, e che riguarda i limiti dei linguaggi elementari. Come si è visto considerando la struttura dei naturali, o quella stessa degli insiemi, molte delle strutture classiche *non* sono caratterizzabili assiomaticamente mediante assiomi elementari e lo stesso vale per classi (non più singole strutture) studiate dall'algebrista. Si pone naturalmente la domanda: quando una classe di strutture è definibile in termini di assiomi *elementari*?

Nel 1955 Tarski (sul modello della varietà di algebre introdotte da Garrett Birkhoff nel 1936) aveva proposto il concetto di *classe elementare*: una classe di strutture che coincide con la classe dei modelli di una teoria finitamente assiomatizzabile; classe elementare *generalizzata* è l'intersezione di una famiglia arbitraria di classi elementari. Con questa terminologia Tarski poneva il problema generale di trovare un punto di contatto tra la forma sintattica degli assiomi che determinano classi elementari e le proprietà di chiusura di queste ultime rispetto a operazioni algebriche. Nel 1955 Jerzy Łos era riuscito a dimostrare contemporaneamente a Tarski che le classi chiuse rispetto al passaggio alle sottostrutture erano tutte e sole quelle corrispondenti a teorie i cui assiomi erano in forma prenessa puramente universale (ossia con solo quantificatori universali nel prefisso). Risultato duale valeva per la chiusura rispetto alle estensioni e agli assiomi puramente esistenziali. Nel 1959 A. Robinson e indipendentemente Łos e C.C. Chang trovavano un'analoga caratterizzazione per le classi elementari chiuse rispetto a somme su catene di strutture, dimostrando che gli assiomi per queste classi sono del tipo $\forall\exists$ (ossia tutti i quantificatori universali precedono i quantificatori esistenziali). Sempre nel 1959 Roger Lyndon dimostrava che le classi elementari chiuse rispetto al passaggio all'immagine omomorfa sono quelle determinate da assiomi «positivi» (ossia che non contengono simboli di negazione e implicazione). Caratterizzazione analoga Lyndon otteneva anche per le classi chiuse rispetto al prodotto sottodiretto.

Tutti questi risultati culminano si può dire nei lavori di H. Jero-

me Keisler che nel 1960 riuscì a fornire un metodo unitario per la dimostrazione di tutti questi teoremi sulla base di una generalizzazione del concetto di diagramma (fondato su una estensione del concetto di formula atomica). Keisler riusciva anche a risolvere il problema della caratterizzazione delle classi elementari determinate da assiomi con prefisso arbitrario e per quelle chiuse rispetto a limiti diretti e inversi e prodotti ridotti. Solo nel 1965 finalmente veniva risolto il problema della caratterizzazione per il prodotto diretto, ad opera di Joseph M. Weinstein.

Rimaneva però aperto un problema più generale: quello di caratterizzare le classi elementari nell'ambito delle classi di strutture. È questo problema che riporta alla questione della caratterizzazione dell'equivalenza elementare: carattere specifico di tali classi (come del loro complemento) è quello di essere chiuse rispetto all'equivalenza elementare. È ancora dovuto a Keisler (1961) il risultato centrale a questo riguardo, fondato sul concetto di *ultraprodotto* introdotto da Łos nel 1955 ed esteso ai sistemi relazionali da Thomas E. Frayne, Ann C. Morel e Dana Scott nel 1961-62. Data una famiglia $\{\mathfrak{A}_i\}_{i\in I}$ di strutture ed un ultrafiltro F nell'algebra delle parti $\mathscr{P}(I)$ di I, l'ultraprodotto associato $\Pi_{i\in I}\mathfrak{A}_i/F$ si ottiene prendendo come dominio l'insieme delle classi di equivalenza del prodotto diretto $\Pi_{i\in I}\, A_i$ rispetto all'equivalenza

$$f\sim g \text{ se e solo se } \{i \in I \mid f(i) = g(i)\} \in F$$

e definendo l'interpretazione dei simboli di relazione e di funzione in modo che per ogni formula atomica $\mathscr{A}$ ed elementi $\hat{f}_1,\dots,\hat{f}_n$ (classi di equivalenza di $f_1,\dots,f_n$) si abbia che

$$\mathfrak{A}_i/F \models \mathscr{A}(\hat{f}_1,\dots,\hat{f}_n) \text{ se e solo se } \{i \in I \mid \mathfrak{A}_i \models \mathscr{A}(f_1(i),\dots,f_n(i))\} \in F.$$

Definiti originariamente da Łos in termini della nozione di *champ logique* come particolari strutture polivalenti, gli ultraprodotti furono presentati come casi particolari di prodotti ridotti (strutture definite come sopra prendendo come F un filtro qualunque, non necessariamente massimale) da Frayne, Morel e Scott in un fondamentale articolo del 1962, dove veniva ottenuta tutta una serie di risultati sui prodotti ridotti che generalizzavano ed estendevano proprietà dei prodotti diretti (i quali ultimi si possono ve-

dere come prodotti ridotti in cui l'ultrafiltro ha come elemento il solo insieme I).

Il teorema fondamentale sugli ultraprodotti dimostrato da Łos, stabiliva che una formula $\mathscr{A}$ è vera nell'ultraprodotto se essa è vera su una famiglia «abbastanza grande» di indici, dove «abbastanza grande» significa appartenente all'ultrafiltro, così come è posto per definizione per le formule atomiche. Da questa caratterizzazione Frayne, Morel e Scott ottenevano immediatamente una dimostrazione del teorema di compattezza, ma i risultati più sorprendenti furono ottenuti negli stessi anni da Jerome Keisler e Simon Kochen in connessione appunto col problema dell'equivalenza elementare. Caso particolare dell'ultraprodotto è infatti l'ultrapotenza, che si ottiene quando tutte le strutture della famiglia sono uguali. Il teorema di Łos garantiva che le ultrapotenze sono estensioni elementari delle strutture di partenza e Keisler riusciva a dimostrare all'inverso che due strutture $\mathfrak{A}$ e $\mathfrak{B}$ sono elementarmente equivalenti se hanno ultrapotenze isomorfe, fornendo così la prima caratterizzazione puramente *algebrica* dell'equivalenza elementare. Da questo risultato discende subito che: una classe K di strutture è elementare (nel senso generalizzato) se 1) K è chiusa rispetto alla formazione di ultraprodotti e passaggio alle immagini isomorfe e 2) $\overline{K}$ è chiusa rispetto al passaggio alle ultrapotenze. Il risultato di Keisler, almeno nella dimostrazione da lui stesso presentata, dipendeva dall'ipotesi del continuo;[2] nello stesso anno però Simon B. Kochen, introducendo il concetto di *ultralimite* (ottenuto da quello di ultraprodotto), riusciva a trovare un'analoga caratterizzazione senza tale ipotesi. Il concetto di ultraprodotto si è rivelato come uno strumento potentissimo per la *costruzione* di modelli e per la sistematizzazione e unificazione di quasi tutta la teoria. In questa direzione l'unico concetto «rivale» è quello di *modello speciale* introdotto da Michael D. Morley e Robert L. Vaught nel 1962 ricollegandosi a precedenti indagini di Jónsson e Fräissé sulle strutture universali e omogenee: si tratta di un tipo particolare di struttura che gode per così dire della proprietà di «rappresentare» classi elementari in forza di una loro «saturazione» e mediante il cui impiego quindi l'articolazione di tutto il discorso risulta

[2] Nel 1971 Saharon Shelah ha potuto eliminare tale ipotesi dalla dimostrazione del teorema.

estremamente elegante e unitaria. Non deve quindi stupire che i due concetti di ultraprodotto e di modello speciale risultino strettamente legati; come infatti ha dimostrato Keisler è possibile costruire mediante ultraprodotti particolari i modelli speciali di cui il teorema fondamentale di Morley-Vaught dimostra l'esistenza ma non fornisce un procedimento di costruzione (si noti d'altra parte che, inversamente, il teorema di Morley-Vaught, formulato in termini di ultraprodotti, ha come corollario il teorema di Keisler).

Modelli speciali e ultraprodotti si sono rivelati estremamente utili anche per affrontare un altro tipo di problemi, quelli della *definibilità*, che sono centrali in ogni studio delle teorie assiomatiche. Ogni teoria è costruita su un linguaggio con uno specifico vocabolario e con specifici assiomi: *teoremi* sono le formule la cui *verità* si riconduce a quella degli assiomi via dimostrazioni, costanti (o concetti) *definibili* entro la teoria sono quelli il cui *significato* si riconduce a quello delle costanti del vocabolario via definizioni. Così, ad esempio, diremo che una costante relazionale P è *definibile* in $\mathfrak{T}$ a partire dalle costanti $Q_1,...,Q_k$ (non tutte necessariamente relazionali) se esiste una formula $\mathscr{A}(x_1,...,x_n)$ che contiene *solo* occorrenze delle costanti $Q_1,...,Q_k$ per cui

$$\text{D} \qquad \forall x_1, ...,x_n(P(x_1, ...,x_n) \longleftrightarrow \mathscr{A}(x_1, ...,x_n))$$

è teorema (o conseguenza) di $\mathfrak{T}$. Un discorso analogo vale per le costanti individuali e quelle funzionali. In tutti i casi, D è una *definizione esplicita* di P a partire da $Q_1,...,Q_k$, e $\mathscr{A}$ è il suo *definiens*.

Quando si assiomatizza una teoria intuitiva (ad esempio la geometria elementare) è estremamente interessante indagare quali concetti si possono assumere come primitivi e quali definire. Come già Peano aveva mostrato all'inizio del secolo, non è sempre possibile aggiungere ad una teoria $\mathfrak{T}$ una formula come la D semplicemente assumendo che sia una definizione che non allarga le capacità deduttive della teoria; questo va provato di volta in volta ed è la proprietà di *non creatività* delle definizioni che fa sì che esse siano pure stipulazioni linguistiche e non assunzioni fattuali aggiuntive. Allo stesso modo – ovviamente – è facile convincersi che la definibilità di un concetto dipende dalla forza degli assiomi scelti, in quanto sono essi che fissano il significato delle costanti primitive.

Problemi di questo tipo avevano svolto un ruolo centrale nelle indagini assiomatiche all'inizio del secolo, come si ricorderà dal paragrafo III,5, in particolare nelle ricerche sui fondamenti della geometria. Tanto in Italia con Peano, M. Pieri, A. Padoa, quanto negli Stati Uniti con O. Veblen, H. Moore, E.V. Huntington, per fare solo i nomi più importanti, molte furono le ricerche su quali primitivi fossero possibili per la geometria del piano e dello spazio, e come queste scelte dipendessero dal tipo di geometria in esame.

È in questo contesto che nel 1901 Padoa aveva formulato un criterio per la definibilità di una costante predicativa da dati assiomi. Il criterio afferma sostanzialmente che se P è definibile da $Q_1, ..., Q_k$ in $\mathfrak{T}$, allora, preso un qualsiasi modello $\mathfrak{M}$ di $\mathfrak{T}$ con dominio D e una qualsiasi biiezione $\varphi: D \rightarrow D$ che conserva il significato di $Q_1, ..., Q_k$, cioè che è un *automorfismo* rispetto ad esse nel senso che per ogni $a_1, ..., a_n$ D se $Q_j^{\mathfrak{M}}$ è l'interpretazione di Q_j in $\mathfrak{M}$

$$<a_1, ..., a_n> \in Q_j^{\mathfrak{M}} \text{ se } <\varphi a_1, ..., \varphi a_n> \in Q_j^{\mathfrak{M}},$$

allora φ è un automorfismo anche rispetto a P. In questo modo si ha un criterio per la *non* definibilità di P del tutto analogo a quello dei contromodelli per la non dimostrabilità: P non è definibile se si trova un modello $\mathfrak{M}$ di $\mathfrak{T}$ e un automorfismo φ rispetto a $Q_1, ..., Q_k$ che non conserva il significato di P in $\mathfrak{M}$. Come tale, il principio proposto da Padoa non solo offriva un contesto unitario in cui situare le ricerche di studiosi come Pieri, Huntington, ecc., che di problemi di definibilità si erano occupati all'inizio del secolo, ma offriva uno strumento concreto che nelle mani di Tarski e Evert Willem Beth (1908-64) negli anni cinquanta permise ad esempio di risolvere in modo esemplare problemi ancora aperti sui concetti primitivi della geometria elementare.

Nel 1959 Tarski aveva pubblicato un lavoro ormai classico *What is elementary Geometry?* (*Cos'è la geometria elementare?*) in cui, riprendendo precedenti ricerche di cui già ci siamo occupati, forniva una assiomatizzazione elementare (cioè al primo ordine) della geometria euclidea. Abbiamo già detto sopra dei risultati di Tarski sulla completezza e decidibilità del sistema: quelli che ancora rimanevano aperti erano problemi sulla definibilità. Sfruttando precedenti ricerche di Pieri, Tarski aveva assunto come primitive

– come si ricorderà – le relazioni *stare tra* e la *equidistanza* tra coppie di punti, in termini delle quali era possibile definire le relazioni fondamentali della geometria del piano. Era naturale chiedersi se si potevano ridurre i primitivi. Nel 1905 Veblen, usando mezzi che trascendevano le capacità espressive dei linguaggi elementari, ritenne di aver mostrato come l'equidistanza (e quindi la congruenza fra segmenti) fosse definibile via un *detour* nella geometria proiettiva a partire dalla relazione – non metrica – *stare tra*, mentre Pieri aveva provato nel 1908 che nello spazio era sufficiente una relazione ternaria tra punti R, per cui $R\ (a,\ b,\ c)$ se e solo se $(\overline{a,b}) = (\overline{b,c})$, cioè a e c giacciono sulla sfera di centro b.

In una serie di articoli della seconda metà degli anni cinquanta, Tarski e Beth riprendevano i lavori che Tarski aveva condotto assieme a Lindenbaum e provavano non solo che il risultato di Veblen era falso (e anzi che col metodo di Padoa si poteva provare il contrario, cioè che *stare tra* non può definire l'equidistanza di coppie di punti) ma pure che il risultato di Pieri non si estendeva al piano, formulando un teorema generale sulle relazioni ternarie e la definibilità degli enti geometrici fondamentali al variare delle dimensioni dello spazio ambiente.

L'utilità pratica del metodo di Padoa (di cui quelle di sopra non sono che alcune applicazioni) portava naturalmente alla domanda se esso fosse *sempre* utilizzabile, cioè se fosse non solo condizione necessaria per la definibilità, ma anche *sufficiente*. Già nel 1936 Tarski e Lindenbaum avevano mostrato che la condizione sugli automorfismi implicava la definibilità una volta che ci si muovesse in un linguaggio d'ordine superiore, in cui si può quantificare su proprietà, relazioni, ecc. Il problema inizia quando si impongono limiti al linguaggio e si richiede ad esempio la definibilità *elementare*. Fu Beth a mostrare che questo era possibile, in un lavoro del 1953 in cui, in un modo piuttosto complicato, dimostrava il converso del principio di Padoa per linguaggi elementari. Nel 1957 W. Craig ridimostrava il teorema – sempre per via sintattica – utilizzando questa volta una formulazione dei calcoli di Gentzen di deduzione naturale partendo dal suo *teorema di interpolazione* che afferma che se $\mathscr{A}$ e $\mathscr{B}$ sono due enunciati tali che

$$\vdash \mathscr{A} \rightarrow \mathscr{B}$$

allora esiste un enunciato $\mathscr{C}$ il cui vocabolario extralogico è quello comune ad $\mathscr{A}$ e $\mathscr{B}$, che interpola l'implicazione, nel senso che

$$\vdash \mathscr{A} \rightarrow \mathscr{C} \text{ e } \vdash \mathscr{C} \rightarrow \mathscr{B}.$$

La prima dimostrazione genuinamente semantica del teorema di Beth venne data da A. Robinson nel 1956, collegandola questa volta ad un teorema di coerenza (oggi noto come *teorema di coerenza congiunta* di Robinson) secondo il quale se $\mathfrak{T}$ e $\mathfrak{T}'$ sono teorie la cui intersezione $\mathfrak{T} \cap \mathfrak{T}'$ è completa nell'intersezione dei linguaggi L e L' di $\mathfrak{T}$ e $\mathfrak{T}'$, allora $\mathfrak{T} \cup \mathfrak{T}'$ è coerente: in altre parole se $\mathfrak{T} \cup \mathfrak{T}'$ permette di derivare una contraddizione, questa è già localizzabile nel linguaggio comune alle due teorie. Dal teorema di coerenza di Robinson segue quello di interpolazione di Craig e il viceversa vale sfruttando essenzialmente il fatto che per i linguaggi elementari si ha il teorema di compattezza. Entrambi i teoremi forniscono così una via d'accesso al teorema di Beth, una più esplicitamente in termini di modelli, l'altra strettamente collegata alle tecniche – illustrate per la prima volta dal ricordato teorema di Henkin – che si basano sulla costruzione di modelli a partire da dati linguistici. In questo senso se la via che passa per il teorema di Robinson fu seguita utilizzando metodi ancor più sofisticati (quali ad esempio i modelli speciali o – come mostrato da Barwise e Schilpf nel 1975 – i modelli *ricorsivamente saturi*), il teorema di Craig e i suoi raffinamenti sono stati al centro di tutta una serie di sviluppi della logica elementare e non elementare in cui si simulano semanticamente le tecniche di Gentzen fornendo metodi *analitici* (che ricorrono cioè alle sole sottoformule degli insiemi dati, come avviene per Gentzen sfruttando l'*erweiterter Hauptsatz*) per costruire modelli.

È questa idea che è al centro delle *tavole* introdotte da Beth nel 1956, proprio proseguendo il tipo di indagini che l'avevano portato a dimostrare il teorema sulla definibilità e a studiare la convergenza di metodi sintattici e semantici. Dato il problema di sapere se $\mathscr{A} \vdash \mathscr{B}$, Beth dava delle regole sistematiche che, partendo dalla ipotesi che $\mathscr{A} \vdash \mathscr{B}$ *non* valga, assegnano valori di verità alle sottoformule di $\mathscr{A}$ e $\mathscr{B}$. Le possibili assegnazioni di verità, che nel nostro caso sono compatibili con l'ipotesi d'assurdo che $\mathscr{A}$ sia vero e $\mathscr{B}$ falso, si dispongono ad albero ed ogni ramo opportunamente completato se non contiene condizioni contraddittorie (come il

fatto che una formula sia tanto vera che falsa) dà origine a un'interpretazione che falsifica il fatto che $\mathscr{A} \vdash \mathscr{B}$. Poiché la procedura è *sistematica* – come si può dimostrare – se non esistono rami non contraddittori $\mathscr{A} \vdash \mathscr{B}$ non è falsificabile e risulta così verificato.

Possiamo illustrare l'uso delle tavole, con un esempio che diamo utilizzando le formule *segnate* introdotte da R. Smullyan nel 1963 per evitare le doppie colonne che Beth impiegava ma che oscuravano la struttura geometrica della dimostrazione. Scriveremo cioè $F\mathscr{A}$ o $V\mathscr{A}$ a seconda che si assuma la falsità o la verità di $\mathscr{A}$. Così, per la formula

$$(\mathscr{A} \to (\mathscr{B} \to \mathscr{C})) \to ((\mathscr{A} \to \mathscr{B}) \to (\mathscr{A} \to \mathscr{C})$$

avremo ad esempio l'albero

$$F((\mathscr{A} \to (\mathscr{B} \to \mathscr{C})) \to ((\mathscr{A} \to \mathscr{B}) \to (\mathscr{A} \to \mathscr{C})))$$

$$V(\mathscr{A} \to (\mathscr{B} \to \mathscr{C}))$$

$$F(\mathscr{A} \to \mathscr{B}) \to (\mathscr{A} \to \mathscr{C})$$

$$V(\mathscr{A} \to \mathscr{B})$$

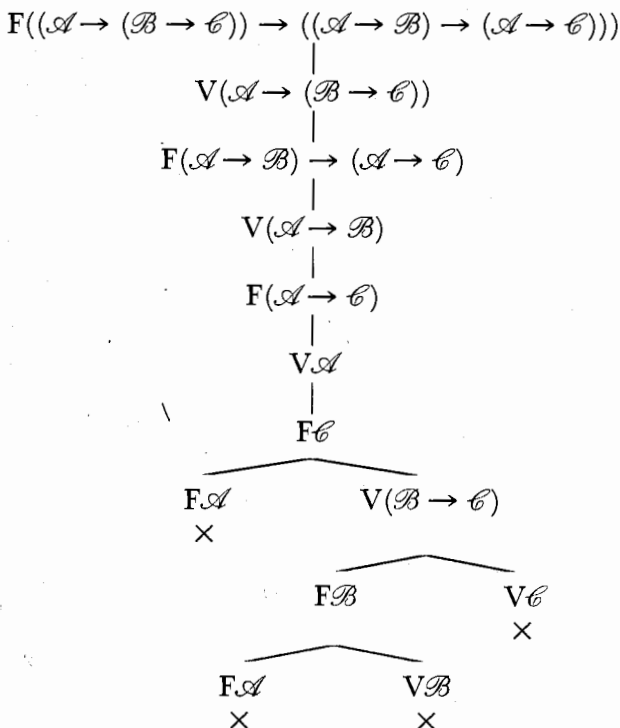
$$F(\mathscr{A} \to \mathscr{C})$$

$$V\mathscr{A}$$

$$F\mathscr{C}$$

$F\mathscr{A}$ × $V(\mathscr{B} \to \mathscr{C})$

$F\mathscr{B}$ $V\mathscr{C}$ ×

$F\mathscr{A}$ × $V\mathscr{B}$ ×

Il segno × sotto i nodi terminali in basso significa che il ramo che termina con il nodo è *chiuso* nel senso che contiene per almeno una formula $\mathscr{A}$ tanto $V\mathscr{A}$ quanto $F\mathscr{A}$.

Le regole che permettono di passare (eventualmente con dira-

mazioni) da un nodo ai successivi sono presto date (ci limitiamo alle regole proposizionali):

$$1)\quad \frac{V \neg \mathscr{A}}{F\mathscr{A}} \qquad \frac{F \neg \mathscr{A}}{V\mathscr{A}}$$

$$2)\quad \frac{V(\mathscr{A} \wedge \mathscr{B})}{\begin{array}{c} V\mathscr{A} \\ V\mathscr{B} \end{array}} \qquad \frac{F(\mathscr{A} \wedge \mathscr{B})}{F\mathscr{A} \mid F\mathscr{B}}$$

$$3)\quad \frac{V(\mathscr{A} \vee \mathscr{B})}{V\mathscr{A} \mid V\mathscr{B}} \qquad \frac{F(\mathscr{A} \vee \mathscr{B})}{\begin{array}{c} F\mathscr{A} \\ F\mathscr{B} \end{array}}$$

$$4)\quad \frac{V(\mathscr{A} \rightarrow \mathscr{B})}{F\mathscr{A} \mid V\mathscr{B}} \qquad \frac{F(\mathscr{A} \rightarrow \mathscr{B})}{\begin{array}{c} V\mathscr{A} \\ F\mathscr{B} \end{array}}$$

dove il segno | indica che l'albero si dirama. L'idea di considerare le tavole come un metodo completo di dimostrazione si basa sull'osservazione che, data una formula $\mathscr{A}$, se esiste un modello in cui $\mathscr{A}$ non è vera, necessariamente utilizzando le regole si otterrà almeno un ramo in ogni possibile dimostrazione che non chiude: quel ramo che contiene per ogni sottoformula atomica $\mathscr{B}$ di $\mathscr{A}$ falsa nel modello, la formula $F\mathscr{B}$ e per ogni sottoformula $\mathscr{C}$ vera, $V\mathscr{C}$. Che all'inverso ogni formula dimostrata risulti valida scende dal fatto che le regole riflettono chiaramente in modo diretto il significato semantico dei connettivi ed è proprio questa che conferisce il carattere *analitico* del calcolo, in quanto le regole fanno passare da formule a sottoformule immediate. Come si vede è in un certo senso la ripresa di quelle tecniche che Skolem aveva introdotto nel 1922 con il suo «calcolo» che si basava sull'idea della costruzione passo passo di contromodelli, solo che mentre Skolem non separava i due momenti della costruzione e della dimostrazione, qui la cosa è resa esplicita.

I vari rami costituiscono i diversi tentativi di costruire contromodelli per la formula data. Che *tutti* i possibili contromodelli siano ottenibili in questo modo è un fatto che si deve dimostrare e che ci dà immediatamente la completezza del calcolo. Nella estensione ai linguaggi del primo ordine, la dimostrazione di

completezza sfrutta – come il lettore si aspetterà – il lemma d König, il cui ruolo qui, come negli esempi che abbiamo già visto, è quello di assicurare l'esistenza di un ramo infinito: il contromodello cercato.

Questa tecnica dei contromodelli riuniva in sé i vantaggi dei metodi semantici e dei vari metodi sintattici (calcoli di Gentzen *in primis*) mostrandone l'unità di fondo: come Smullyan ha sottolineato, le regole del calcolo **LK** di Gentzen non sono che la versione «capovolta» di quelle di Beth, in cui l'antecedente della regola diventa il conseguente e viceversa. Non stupisce che queste tecniche si siano rivelate utili quando si è cercato di estendere il metodo della costruzione di modelli a partire da dati linguistici a linguaggi più forti di quelli elementari nei quali non vale la compattezza e non sono disponibili risultati forti come appunto il teorema di Robinson. In molti di questi casi, teoremi di interpolazione come quello di Craig *sono* dimostrabili e il metodo migliore è il ricorso a questi sistemi analitici. È in quest'ottica che, indipendentemente da Beth, K. J. Hintikka e K. Schütte e quindi R. Smullyan e Melvin Fitting, hanno sistematicamente sviluppato un approccio alle varie logiche – classiche e non, elementari e d'ordine superiore, infinitarie e con quantificatori generalizzati – utilizzando i metodi analitici e mostrandone i vari legami con altre tecniche. Soprattutto dopo i lavori di Mikail Makkai del 1969 in cui si mostrava il ruolo di tali metodi nello studio sistematico dei problemi di interpolazione e preservazione per i linguaggi infinitari, la tecnica è divenuta pressoché universale in questo tipo di questioni e non solo ha permesso di ottenere anche per linguaggi più forti di quelli elementari alcuni dei rafforzamenti del teorema di definibilità di Beth dimostrati da Lars Svenonius, Lyndon (che studiò nel 1959 i legami tra formule preservate da omomorfismi e interpolanti positive) Chang e Makkai, ecc. ma ha ampliato il suo raggio d'applicazione ai *teoremi di omissione dei tipi* e al forcing generalizzato, come mostrato in un lavoro fondamentale di Keisler del 1973. Di alcuni di questi lavori avremo occasione di parlare più avanti. Volgiamoci ora ad un altro dei grandi temi della ricerca in teoria dei modelli.

Abbiamo parlato ripetutamente di linguaggi più forti di quelli elementari e dei limiti di questi ultimi. È un tema di estrema importanza che nel periodo che stiamo studiando acquistò un grande significato.

La caratterizzazione delle classi elementari data sopra dimostra come i linguaggi elementari non siano spesso in grado di discriminare a sufficienza tra modelli non isomorfi, come del resto già mostrava il teorema di Löwenheim-Skolem. L'introduzione del concetto di estensione elementare aveva permesso di cogliere più appieno nel suo significato questa limitazione e di estendere il risultato originale di Skolem nelle due diverse direzioni descritte nella nota 1 a pag. 553. È quindi chiaro che nessuna teoria con un modello infinito, se espressa nel linguaggio elementare, può essere categorica. È per questa ragione che Łos introdusse nel 1954 il concetto più debole di *categoricità in potenza* (da noi già ricordato a proposito del criterio di Vaught) che in effetti acquistò il suo pieno significato dopo che nel 1963 Morley dimostrò un teorema fondamentale secondo il quale se una teoria $\mathfrak{T}$ è categorica in una potenza $\aleph_\alpha > \aleph_0$, allora è categorica in *ogni* potenza non numerabile (si badi bene che ciò *non* significa che la teoria in questione sia *tout court* categorica). Il risultato di Morley – particolarmente interessante per l'enorme serie di collegamenti e sviluppi cui ha dato luogo e di cui avremo occasione di riparlare – ci permette di concludere che per una teoria elementare quattro sono le possibilità per quanto riguarda la categoricità: o è categorica in *ogni* cardinale (come la teoria dell'identità) o è categorica in $\aleph_0$ ma in *nessun* cardinale più che numerabile (come la teoria degli ordini densi senza estremi) o, all'inverso, è categorica in *ogni* cardinale più che numerabile ma non in $\aleph_0$ (come la teoria dei campi algebricamente chiusi di caratteristica zero) o, infine, non è categorica in *nessun* cardinale (come ad esempio la logica pura). In conclusione, la discriminante (per quanto riguarda i cardinali infiniti) è tra $\aleph_0$ e più che numerabile: tutti i cardinali più che numerabili sono solidali fra loro.

Con questo risultato, Morley mostrava i limiti delle possibilità espressive dei linguaggi elementari e rendeva, se così si può dire, del tutto «naturali» quei tentativi iniziati attorno al 1957 da parte di Tarski, Henkin, Scott, Mostowski e altri di ampliarne variamente la potenza espressiva introducendo vari linguaggi rafforzati, che sono oggi al centro dell'attenzione per quanto riguarda la teoria dei modelli e la teoria della dimostrazione.[3] Si tratta di

[3] Il lettore si chiederà senza dubbio come mai, una volta accertati i «limiti» posti abbondantemente in evidenza per i linguaggi elementari, non ci si risolva a passare direttamente a studiare le teorie espresse nel più forte e apparentemente

estensioni di tipo diverso e accenneremo soltanto ad alcune di esse, che hanno permesso di ottenere risultati interessanti, ma ci soffermeremo un po' più a lungo su quei linguaggi che riteniamo tra i più significativi, e sono noti come *linguaggi infinitari*. Tra le estensioni del primo tipo ricordiamo le cosiddette logiche del secondo ordine debole,[4] introdotte da Tarski nel 1958 e che consistono nell'aggiungere a un generico linguaggio elementare $\mathscr{L}$ quantificatori applicabili non a predicati arbitrari (come avverrebbe al secondo ordine) bensì solo a predicati che vengano interpretati su insiemi *finiti*; oppure linguaggi con *quantificatori generalizzati* introdotti da Mostowski nel 1957 nei quali a $\mathscr{L}$ vengono aggiunti quantificatori Q_α in modo tale che un'espressione del tipo $Q_\alpha x \mathscr{A}(x)$ venga interpretata come «esistono almeno $\aleph_\alpha$ individui che soddisfano $\mathscr{A}(x)$». I linguaggi con i quantificatori generalizzati sono tornati al centro dell'attenzione nell'ultimo decennio, ma negli anni sessanta furono i linguaggi infinitari la novità più cospicua.

I linguaggi infinitari introdotti da Tarski e Scott attorno al 1957 e sistemati in modo organico su base algebrica nel 1964 da Carol R. Karp si fondano su un'altra prospettiva, più esplicitamente sintattica. Si tratta cioè di ammettere nel linguaggio espressioni

adeguato linguaggio del secondo ordine. Il fatto è che esistono almeno due buone ragioni, strettamente legate l'una all'altra, che sconsigliano – se addirittura non rendono impossibile – tale passaggio: in primo luogo l'incompletezza semantica dimostrata da Gödel per teorie espresse in un linguaggio di questo tipo, rispetto alla normale accezione di modello, che taglia alle radici ogni possibilità di fare uso di un teorema centrale della teoria dei modelli, il teorema di compattezza, con la conseguente impossibilità di organizzare una soddisfacente teoria dei modelli per tali linguaggi; in secondo luogo, la difficoltà di stabilire una connessione rigorosamente dominabile tra modelli di teorie formulate al primo ordine e le stesse teorie formulate al secondo. Lo spazio tra primo e secondo ordine linguistico si rivela quindi estremamente ampio e l'introduzione dei linguaggi intermedi cui si accenna nel testo, basati come sono sull'approssimazione alla quantificazione del secondo ordine, ha da parte sua il vantaggio di contemperare la maggior potenza espressiva con la possibilità di trasportare ad essi buona parte della concettualizzazione fondamentale della semantica del primo ordine. Criterio generale nell'introduzione di questi linguaggi è quindi la «trasportabilità» ad essi dei teoremi semantici più significativi, in primo luogo del teorema di completezza. È questo un tema su cui torneremo più avanti parlando delle ricerche di Lindström.

[4] Il maggior interesse, forse, del secondo ordine debole sta, come ha mostrato Montague nel 1971, nel fatto che data una teoria $\mathfrak{T}$ del primo ordine, è possibile ottenere in modo canonico una sua estensione al secondo ordine debole che esprime gran parte della semantica di $\mathfrak{T}$.

infinitamente lunghe riprendendo la pratica di Schröder ed i tentativi di Zermelo, Ramsey ed altri nei primi decenni del secolo; va subito notato che, comunque la cosa venga realizzata, le espressioni linguistiche diventano così oggetti infinitari e quindi già per la trattazione della loro sintassi sono necessarie metateorie più forti quali ad esempio particolari frammenti della teoria degli insiemi. Seguendo la sistemazione della Karp si è oggi soliti indicare tali linguaggi col simbolo generico $L_{\alpha,\beta}$ dove α e β sono due ordinali con $\alpha \geqslant \beta \geqslant \omega$. α indica il limite (*non* raggiungibile) della lunghezza delle congiunzioni o disgiunzioni ammesse nel linguaggio, mentre β rappresenta l'analogo limite per le successioni di variabili vincolate da un singolo quantificatore ammissibile in una formula. Così ad esempio nel linguaggio $L_{\omega_1,\omega}$ sono ammesse congiunzioni e disgiunzioni di lunghezza $< \omega_1$ e successioni di variabili quantificate di lunghezza $<\omega$ (ossia solo finite). Si noti che un normale linguaggio $\mathscr{L}$ del primo ordine si ritrova come caso particolare di questi linguaggi infinitari e coincide precisamente con $L_{\omega,\omega}$. Come abbiamo già accennato alla nota 1 di pagina 553 per i linguaggi $L_{\alpha,\beta}$ ($\alpha > \omega$) si può tradurre (con le opportune modificazioni) il teorema di Löwenheim-Skolem «inferiore» (il che in certo senso dà una misura della «lontananza» di questi linguaggi dal secondo ordine totale); viceversa non è in generale possibile estendere l'analogo teorema «superiore» perché per la maggioranza di questi linguaggi non vale il teorema di compattezza, che abbiamo visto essere essenziale nella dimostrazione dell'*upward*. In ogni caso, come provato da William Hanf nel 1962, per linguaggi le cui formule costituiscano un insieme e non una classe propria, è possibile stabilire alcuni fatti che generalizzano le due versioni del teorema di Löwenheim-Skolem. Per ognuno di questi linguaggi possiamo definire due cardinali, $\mathfrak{h}$ e $\mathfrak{l}$ rispettivamente, il numero di Hanf e il numero di Löwenheim del linguaggio, ponendo che $\mathfrak{h}$ è il minimo cardinale α (se esiste) per cui se un insieme di formule del linguaggio ha un modello di cardinalità almeno α, ha modelli di cardinalità arbitrariamente grande; dualmente $\mathfrak{l}$ è minimo cardinale α (se esiste) tale che se un insieme di enunciati del linguaggio ha un modello, allora ha un modello di cardinalità $\leqslant \alpha$. Hanf – generalizzando il teorema superiore di Löwenheim-Skolem – provava che *ogni* linguaggio del tipo da noi considerato ha un numero di Hanf e lo stesso vale per il numero di Löwenheim, fatto questo

che generalizza la versione inferiore del teorema. Per linguaggi elementari con α costanti relazionali (α infinito) è facile vedere che $\mathfrak{h}$ coinciderà con $\aleph_0$ ed $\mathfrak{l}$ con α. Ben più difficile si è dimostrato il problema di computare il numero di Hanf per linguaggi più generali, come testimoniano le numerose ricerche al riguardo.

Ben più complessa la questione per quel che riguarda la completezza (semantica) di tutte queste estensioni dei linguaggi elementari, questione di particolare importanza, come detto sopra, in quanto vincola la dominabilità, dal punto di vista della teoria dei modelli, dei linguaggi stessi. I teoremi di completezza, vincolando la nozione di formula valida a quella di «generabilità mediante un sistema di assiomi» (la «dimostrabilità»), possono essere sostanzialmente intesi come affermanti il fatto che le formule valide costituiscono un insieme in qualche modo «dominabile»: così ad esempio per $\mathscr{L}$ tale insieme risulta essere ricorsivamente enumerabile. Nelle estensioni sopra ricordate, i corrispondenti insiemi di formule valide risultano difficilmente localizzabili dal punto di vista della loro effettività, nel senso che o hanno localizzazioni «inattese», o addirittura sfuggono a tutti gli immediati criteri di classificazione: così ad esempio per $\mathscr{L} + Q_0$ (l'estensione di $\mathscr{L}$ ottenuta aggiungendo il quantificatore «esistono almeno $\aleph_0$...») si ha che l'insieme delle formule valide è analitico (della forma Π^1_1) ma non iperaritmetico; per $\mathscr{L} + Q_1$, viceversa, Vaught ha dimostrato che il corrispondente insieme è ancora ricorsivamente enumerabile; per il linguaggio del secondo ordine si sa solo che l'insieme delle formule valide è non analitico (ed è aperto il problema se sia costruibile nel senso di Gödel) mentre per i linguaggi $L_{\alpha,\beta}$ in generale non si dispone neppure di opportune gerarchie che permettano di esprimere un analogo del teorema di completezza, a meno di ricorrere a quelle generalizzazioni della teoria della ricorsività di cui si è già detto a suo tempo e il cui ruolo in questi contesti è stato sottolineato dai lavori di C. Karp, R. Jensen, A. Kino, G. Takeuti, ecc.

Per il linguaggio $L_{\omega_1,\omega}$ invece la Karp ha dimostrato nel 1964 un opportuno teorema di completezza, ma si può provare che per $L_{\omega_1,\omega}$ non vale il teorema di compattezza ed è questo fatto che impone alla teoria dei modelli per questo linguaggio un'impostazione e un andamento «nuovi» che le conferiscono un particolare e genuino interesse. Si riesce infatti a ridimostrare per questo lin-

guaggio una serie di risultati analoghi a quelli di Łos, Robinson ecc. ricorrendo però, in mancanza della compattezza, a risultati *sintattici*, in particolare all'*Hauptsatz* di Gentzen esteso a questi linguaggi nel 1964 per la prima volta da E.G.K. Lopez-Escobar. In questo modo la teoria della dimostrazione per i linguaggi infinitari sembrerebbe mostrarsi strumento d'analisi molto più sottile della teoria dei modelli. Che ciò non sia vero risulta dallo sviluppo di quelle tecniche analitiche di cui abbiamo parlato, mediante le quali è possibile dare una traduzione più esplicitamente semantica di tutti questi concetti sulla base della nozione di *proprietà generale di coerenza.*

Questi metodi si applicano anche al caso di un tipo più «sottile» di logiche infinitarie, non basato su una considerazione puramente cardinale della «lunghezza» delle formule, ma sulle loro caratteristiche dal punto di vista dell'effettività. Abbiamo già accennato, parlando della ricorsività, ai legami che esistono fra il concetto di finito e quello di effettivo; basandosi su questo fatto, nel 1968 Kreisel, dopo aver notato come espressioni infinitarie fossero intervenute in modo «naturale» già allo stesso inizio del discorso formalizzante ad esempio in Löwenheim, Skolem ed esplicitamente nel 1935 in Zermelo, proponeva di considerare formule di lunghezza «$\mathfrak{A}$-finita» e regole di deduzione che consentissero derivazioni «$\mathfrak{A}$-finite», dove il concetto di $\mathfrak{A}$-finitezza costituisce una generalizzazione infinitaria del concetto di finito, basata su un'analogia – non cardinale ovviamente – ma riguardante la forma di definibilità degli insiemi $\mathfrak{A}$-finiti. In questo modo Kreisel si ricollegava ai tentativi di generalizzazione della teoria della ricorsività, ai cosiddetti ordinali *ammissibili* (introdotti da Saul A. Kripke nel 1964). Sulla scorta delle osservazioni di Kreisel, J. Barwise introduceva i nuovi linguaggi L_A come «frammenti» del più forte linguaggio $L_{\omega_1,\omega}$ dove A è un insieme «ammissibile» come gli insiemi della forma L_α per α ordinale ammissibile; per i linguaggi L_A si riusciva non solo a dimostrare la completezza, ma addirittura una forma di compattezza, mostrando così come la generalizzazione proposta da Kreisel fosse estremamente più «fine» di quella basata su considerazioni puramente cardinali.

Il fatto che si sia ritenuto opportuno, per definire in modo più adeguato proprietà di strutture, estendere il linguaggio $\mathscr{L}$ (pur senza giungere al secondo ordine) non deve indurre il lettore a

credere che i linguaggi elementari siano scarsamente utili e quindi ormai abbandonati nelle indagini di teoria dei modelli. Tutto un compatto *corpus* di ricerche strettamente collegate alla concreta pratica matematica è stato sviluppato a partire dalla fine degli anni sessanta sfruttando proprio i «modelli indesiderati», non standard, che le teorie del primo ordine ammettono. Al di là dei risultati «locali» ottenuti da Skolem cui abbiamo accennato a suo tempo, e delle immediate conseguenze a questo riguardo del teorema di incompletezza di Gödel, a partire dal 1960 A. Robinson ha sviluppato un approccio alla matematica «non standard» il quale – oltre al suo intrinseco interesse dal punto di vista logico e matematico – ha una specifica portata proprio dal punto di vista storico, in quanto in particolare i sistemi di Analisi non standard, presentati da Robinson nel 1966, recuperano pienamente le impostazioni e intuizioni leibniziane circa un'Analisi nella quale avesse normale diritto di cittadinanza il concetto di infinitesimo.

A parte però quest'aspetto, va ribadito che l'Analisi non standard di Robinson, come vedremo più ampiamente in VII, 3, ha permesso non solo di riottenere in modo elegante significativi risultati di Analisi classica, ma ha addirittura consentito di risolvere nuovi problemi. L'essenziale è che gli sviluppi più significativi di queste ricerche si situano in un contesto di linguaggi di ordine superiore o utilizzando il linguaggio della teoria degli insiemi, sfruttando in modo positivo l'apparente limite costituito dall'assenza di uno strumento così forte ed essenziale come è quello rappresentato dal teorema di completezza e compattezza rispetto ai modelli normali.

Le estensioni non standard sono appunto modelli non normali (se lo fossero non potrebbero avere elementi infinitesimi, in quanto la non esistenza di questi ultimi segue dall'assioma di Archimede che è formulabile al secondo ordine pieno (ristretto a modelli normali) ma non entro linguaggi elementari). Le ricerche in questo campo sono state estremamente attive e approfondite e hanno messo in luce come risultati apparentemente limitativi non costituiscano un banale e bloccante «incidente sul lavoro» ma un potente strumento di ulteriore analisi (e viene naturale l'associazione col ruolo svolto, in questo senso, dalle antinomie, pur se in contesto completamente diverso). Sulla matematica non standard torneremo comunque più avanti.

Anche se finora il nucleo del discorso è stato rivolto all'aspetto *semantico* della logica, abbiamo avuto più volte l'occasione di fare riferimento all'altro approccio allo studio delle teorie, proprio della teoria della dimostrazione. Il fatto non è casuale in quanto un aspetto importante della moderna teoria della dimostrazione è appunto il superamento dell'originario programma hilbertiano mediante il ritorno a un'analisi dei concetti intuitivi fondamentali lungo una via che già abbiamo vista aperta – come precursore – da Gentzen. Non deve quindi stupire che in questa ricerca della propria autonomia e ragion d'essere (che cosa differenzia un'analisi delle teorie dal punto di vista sintattico rispetto a un'analoga indagine di tipo semantico?) la teoria della dimostrazione a partire dagli anni cinquanta non si presenti più (cosa che avveniva per la *Beweistheorie* hilbertiana) come tentativo di *ridurre* le teorie formalizzate ad una forma dominabile, assunta sin dall'inizio e una volta per tutte come elementare, mediante strumenti limitati e precisati a priori, bensì estenda il suo campo di indagine più direttamente al concetto generale di dimostrazione, alle sue classificazioni in base ai gradi di evidenza, portando ad un ripensamento di quegli stessi concetti di «finito», «costruttivo» ecc. che per Hilbert erano il dato di partenza inquestionabile piuttosto che il punto di arrivo. È a Georg Kreisel che si deve principalmente la ripresa e la nuova impostazione del programma hilbertiano.

Nel 1958 in un articolo fondamentale intitolato *Hilbert's programme* (*Il programma di Hilbert*) Kreisel sottolineava: 1) come Hilbert non avesse mai condotto un'analisi precisa del concetto di matematica finita lasciando così sostanzialmente nel generico tutto il suo programma; 2) come la limitazione al solo problema della dimostrazione di coerenza fosse arbitraria e ingiustificata. Occorreva, per far fronte al teorema di Gödel, procedere lungo una doppia direzione: da una parte, studiando il meccanismo di riduzione mediante il quale attraverso l'analisi delle sue dimostrazioni una teoria viene interpretata all'interno di un'altra teoria (nel caso di Hilbert: la matematica finitista), si doveva giungere a una classificazione, in base alla loro «evidenza», di queste teorie fondanti; dall'altra, era necessario analizzare in generale il concetto intuiti-

vo di dimostrazione, saggiandone le varie precisazioni formali senza porsi limiti riduttivi *a priori* sui metodi ammessi in questa analisi.

Kreisel – e dopo di lui molti altri – ha dato sostanza a questo programma in diversi lavori in cui ha cercato di mostrare come ci sia un aspetto nell'analisi delle dimostrazioni che – al di là di ogni programma riduzionista – ha un significato matematico profondo. È la possibilità di indagare, sia che si applichino le tecniche di Gentzen o di Herbrand o l'ε-calcolo di Hilbert, il significato *costruttivo* di teoremi matematici, cercando di estrarre prescrizioni per computare confini numerici e ottenere in generale informazioni di carattere computazionale. Già negli anni cinquanta Kreisel aveva mostrato come in questa possibilità stesse il significato *matematico* delle dimostrazioni di coerenza fornendo esempi tratti dall'algebra (quali il *teorema degli zeri* di Hilbert e il teorema di Artin sulle somme di quadrati, che A. Robinson aveva analizzato con la sua tecnica della model-completezza) come pure dalla teoria dei numeri e dall'Analisi. Per Kreisel – al di là di ogni atteggiamento ideologico di carattere fondazionale – l'aspetto centrale della metamatematica hilbertiana ed uno degli obiettivi della teoria della dimostrazione doveva essere quello di raffinare gli strumenti d'analisi sì da permettere di indagare dal punto di vista costruttivo la matematica classica ed intuizionista, il grado di evidenza delle diverse dimostrazioni di uno stesso risultato, in generale il significato effettivo di ciò che il matematico prova. Da allora i risultati in questa direzione si sono moltiplicati e si può dire che alcune delle tecniche più raffinate elaborate al riguardo siano collegate a questi obiettivi.

È in questo contesto che l'impostazione di Gentzen si è andata esplicitando in tutta la sua portata a cominciare dalle ricerche di Kurt Schütte dedicate alla teoria dei tipi (1959) fino a quelle di William Tait sui linguaggi e le regole infinitarie. Schütte e Tait riuscirono a rendere più trasparente la concezione generale di Gentzen relativa alla coerenza delle teorie quali l'aritmetica, ponendone in luce in modo più organico e sistematico gli elementi infinitari. Schütte otteneva così una dimostrazione di coerenza dell'aritmetica in cui l'induzione fino a ω viene sostituita da una regola infinitaria – la cosiddetta ω-regola – che può esprimersi come segue:

$$\frac{P(\bar{0}), P(\bar{1}), \ldots, P(\bar{n}), \ldots}{\forall x P(x)}$$

Questa regola aveva una lunga storia che possiamo far risalire ad alcuni interventi di Langford del 1927 e al tentativo di Carnap del 1930 di superare gli ostacoli posti dalle definizioni impredicative, formulando l'induzione come principio d'inferenza.

In direzione analoga Tait affrontava direttamente lo studio di particolari linguaggi infinitari dal punto di vista dei sistemi della deduzione naturale. In questo modo, abbandonando la limitazione ai sistemi formali nel senso di Hilbert via l'ammissione di regole di formazione (Tait) o deduzione (Schütte) infinitarie, la necessità del ricorso all'induzione fino a ε_0 risultava localizzabile con precisione: erano gli alberi infiniti di dimostrazione che una volta opportunamente indiciati da ordinali che ne misurano la complessità imponevano questo passaggio. Dimostrare una proprietà per tutte le dimostrazioni di un dato sistema equivale infatti a imporre l'utilizzazione di un principio di induzione che si sviluppi lungo un «segmento» degli ordinali sufficientemente esteso per «ricoprire» tutti i possibili tipi di dimostrazione (alberi) che occorrono nel sistema dato; così interpretati, i principi di induzione transfinita acquistano un significato ben determinato e si possono allora estendere a sistemi più forti, in grado di formalizzare teorie quali l'Analisi (o suoi sottosistemi). Negli anni sessanta i lavori in questa direzione si moltiplicano e sono dovuti soprattutto oltre a Schütte e Tait in particolare a Takeuti (che propone la tecnica dei cosiddetti *diagrammi ordinali*).

Proprio nel 1958 però, come già accennato, Kurt Gödel presentava un approccio apparentemente molto diverso all'analisi delle teorie, approccio che già aveva sviluppato nel 1941. L'idea di fondo consisteva nell'associare ad ogni formula derivabile di una teoria (in particolare dell'aritmetica $\mathfrak{P}$ di Peano o di quella intuizionista $\mathfrak{AS}$) un'altra formula del tipo $\forall y_1,\ldots,y_k \exists x_1,\ldots,x_n \mathscr{A}$, dove $\mathscr{A}$ non contiene quantificatori e $y_1, \ldots, y_k, x_1, \ldots, x_n$ sono variabili per funzionali ricorsivi di tipo finito (definibili mediante gli ordinari schemi di ricorsione), funzionali cioè che hanno come argomenti e valori ancora funzionali e che sono di tipo finito nel senso della teoria dei tipi. L'associazione di queste formule a quelle del sistema originario avviene seguendo passo passo le derivazioni di que-

ste ultime, e la complessità delle definizioni dei funzionali che vengono introdotti riflette quella delle regole di deduzione applicate nel corso della derivazione stessa. Le formule così associate significano in altre parole che esistono funzionali in grado di «scegliere» gli individui opportuni per soddisfare le quantificazioni della formula tradotta e forniscono una interpretazione del significato costruttivo delle formule aritmetiche. È questa l'*interpretazione Dialectica* – dal nome della rivista in cui comparve l'articolo in questione. Per dimostrare la coerenza del sistema che si analizza occorre quindi poter provare entro la teoria T dei funzionali che viene utilizzata per interpretarne le formule, l'*esistenza* di questi funzionali, cioè la validità dei principi di ricorsione necessari per definirli. Nel caso dell'aritmetica – come ricordato – questi principi di ricorsione sono necessariamente transfiniti in quanto non ricavabili dagli ordinari schemi di ricorsione per *funzioni* dimostrabili in base all'induzione su ω: l'importante è che è sufficiente l'induzione fino a ε_0 per giustificarli. Vale a dire, l'induzione sino a ε_0 ci permette di costruire un modello i cui oggetti sono funzionali che soddisfano gli assiomi della teoria T che risulta così coerente.

L'obiettivo di Gödel era di indagare più da vicino in che senso la nozione intuizionista di dimostrazione è davvero costruttiva e fornire così – tra l'altro – un metodo per controllare quando date formule sono teoremi intuizionisti e quindi anche quando la teoria in oggetto è coerente. Nel suo lavoro Gödel prendeva in considerazione l'aritmetica intuizionista $\mathfrak{AS}$ e analizzava mediante la sua interpretazione in termini di funzionali *questa* teoria, non l'aritmetica di Peano $\mathfrak{P}$. L'interpretazione che Gödel definiva – come Gödel stesso dichiarò esplicitamente a Kreisel – cercava di porre in luce entro che limiti le dimostrazioni esistenziali su base intuizionista siano in grado di fornire realizzazioni esplicite. In forza del teorema di incompletezza e della traduzione dell'aritmetica di Peano in $\mathfrak{AS}$, sappiamo che queste realizzazioni esplicite dei quantificatori esistenziali non potevano in generale essere funzioni numeriche di carattere ricorsivo. Ciò può avvenire – come aveva dimostrato Kleene nel 1945 provando che in $\mathfrak{AS}$ per ogni formula

$$\forall x_1 \ldots x_n \, \forall \exists y \mathscr{A}(x_1, \ldots, x_n, y)$$

dimostrabile esiste un termine t per cui

$$\forall x_1, \ldots, x_n \, \mathscr{A}(x_1, \ldots, x_n, t(x_1, \ldots, x_n))$$

è un teorema – solo per formule di complessità ridotta, ma non per *ogni* formula. Deve quindi comparire un elemento infinitario astratto per poter esemplificare ogni esistenziale.

Questo elemento astratto – ma non per questo non costruttivo – viene localizzato da Gödel nel passaggio ai tipi superiori: occorrono funzioni che hanno come argomenti funzioni, funzioni di funzioni, ecc. L'importante è che questi *funzionali* siano definiti esplicitamente e in modo costruttivo utilizzando schemi di ricorsione a partire da funzioni costruttive. La teoria T in cui Gödel interpreta $\mathfrak{AS}$ è una teoria con variabili per funzionali di ogni tipo finito, le costanti $\overline{0}$ e S (per zero e successore) e per ogni tipo σ un funzionale R_σ, il recursore, che ci permette di costruire i termini che vengono interpretati come funzionali di tipo σ definiti per ricorsione. La teoria è essenzialmente equazionale e non usa quantificatori. Come ricordato, l'interpretazione di Gödel associa ad ogni $\mathscr{A}$ di $\mathfrak{AS}$ una $\mathscr{A}^{\mathrm{D}}$ del tipo

$$\exists y_1 \ldots y_n \forall x_1, \ldots x_k \mathscr{A}_D$$

dove $\mathscr{A}_D$ è una formula di T. Se con $\boldsymbol{U}$, $\boldsymbol{Y}$, indichiamo funzionali di tipo opportuno, con le successioni di variabili $\boldsymbol{x}$, $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{y}$ che supponiamo disgiunte e per ipotesi induttiva assumiamo che

$$\mathscr{A}^D \equiv \exists \boldsymbol{x} \forall \boldsymbol{y} \mathscr{A}_D(\boldsymbol{x}, \boldsymbol{y}),\ \mathscr{B}^{\mathrm{D}} \equiv \exists \boldsymbol{u} \forall \boldsymbol{v} \mathscr{B}_D(\boldsymbol{u}, \boldsymbol{v})$$

avremo

$$\begin{array}{lll}
(\wedge^D) & (\mathscr{A} \wedge \mathscr{B})^D : = & \exists \boldsymbol{xu} \forall \boldsymbol{vy} (\mathscr{A}_D \wedge \mathscr{B}_D), \\
(\vee^D) & (\mathscr{A} \vee \mathscr{B})^D : = & \exists z^0 \boldsymbol{xu} \forall \boldsymbol{vy} [(z = 0 \wedge \mathscr{A}_D) \vee (z = 1 \wedge \mathscr{B}_D)], \\
(\forall^D) & (\forall z \mathscr{A}(z))^D : = & \exists \boldsymbol{x} \forall z \boldsymbol{y} \mathscr{A}_D(\boldsymbol{x}z, \boldsymbol{y}), \\
(\exists^D) & (\exists z \mathscr{A}(z))^D : = & \exists z \boldsymbol{x} \forall \boldsymbol{y} \mathscr{A}_D(z, \boldsymbol{x}, \boldsymbol{y}), \\
(\rightarrow^D) & (\mathscr{A} \rightarrow \mathscr{B})^D : = & \exists \boldsymbol{UY} \forall \boldsymbol{xv} (\mathscr{A}_D(\boldsymbol{x}, \boldsymbol{Yxv}) \rightarrow \mathscr{B}_D(\boldsymbol{Ux}, \boldsymbol{v})).
\end{array}$$

Ciò che Gödel dimostra è che per *ogni* teorema $\mathscr{A}$ di $\mathfrak{AS}$ i funzionali la cui esistenza è affermata da $\mathscr{A}^D$ esistono davvero per T, cioè esistono termini in T per cui si può provare

$$\mathscr{A}^D(t_1, \ldots, t_n, x_1, \ldots, x_k)$$

per ogni $x_1, \ldots, x_k$. La forza di T è data appunto dai principi di definizione per ricorsione che ci permettono di costruire i termini che defi-

niscono gli opportuni funzionali e misurare la complessità dei teoremi di $\mathfrak{AS}$.

L'analisi di Gentzen e quella di Gödel portano quindi allo stesso principio di induzione ma per due vie completamente diverse. Nel primo caso è un principio infinitario che si applica a oggetti concreti (gli alberi di dimostrazioni); nel secondo è uno schema finitario (la ricorsione per la definizione dei funzionali è infatti primitiva) applicato però ad oggetti astratti, i funzionali appunto. È proprio in questo senso che Gödel, parafrasando un'osservazione di Bernays, afferma che «poiché la matematica finitista è definita come quella dell'evidenza *intuitiva*, ciò significa... che per la dimostrazione di non contraddittorietà della teoria dei numeri sono necessari certi concetti *astratti*».

La connessione tra alberi di dimostrazione e funzionali può essere ulteriormente estesa. Nel 1968 Tait seguendo un'indicazione di Curry e Howard è riuscito a sviluppare un collegamento basato su questo semplice fatto: ad ogni derivazione del calcolo di Gentzen si può associare la costruzione di un termine che in un certo senso «definisce» la derivazione stessa; gli operatori che occorrono nel termine sono in corrispondenza con quelli logici e le regole di deduzione con le regole di trasformazione tra termini. Vedremo in seguito come questa idea sia suscettibile di sviluppi profondi, una volta introdotto il discorso sui combinatori che consentono di formulare in astratto in modo esplicito legami tra dimostrazioni, termini che definiscono funzionali e combinatori. Basti per ora osservare che anche i funzionali si possono considerare come denotati da termini e si può quindi stabilire direttamente un collegamento tra i termini costruiti nei due modi. L'elemento comune che risulta da tutti questi molteplici collegamenti è costituito dal fatto che in un modo o nell'altro queste diverse analisi conducono sempre a strutture *fondate* (termini o derivazioni o definizioni di funzionali) la cui complessità può essere studiata in base alla struttura fondata per eccellenza: quella dei numeri ordinali. È proprio questo fatto che spiega l'interesse del tutto particolare che tale struttura ha avuto e continua ad avere in queste ricerche, e che giustifica i numerosi tentativi di individuazione di classi di ordinali che si possano per così dire «nominare» e definire in modo effettivo.

Questo intervento essenziale degli ordinali risulta anche in un altro tipo di indagine riguardante, più che la coerenza, la *comple-*

tezza delle teorie e basato sulla costruzione di successioni transfinite di teorie (indiciate da ordinali *costruttivi*), tali che ogni teoria elemento della successione sia «più forte» della precedente, nel senso che in essa è dimostrabile il cosiddetto *principio di riflessione* per la più debole. I principi di riflessione, studiati sistematicamente per la prima volta da Kreisel e Levy nel 1968, sono schemi dalla forma Dim $\ulcorner\mathscr{A}\urcorner \rightarrow \mathscr{A}$ e sostanzialmente affermano che ogni formula dimostrabile in una data teoria $\mathfrak{S}$ è vera: hanno quindi come ovvio corollario la coerenza del sistema stesso; come infatti dimostravano Kreisel e Levy, $\mathfrak{P}$ rafforzato con una forma dello schema di riflessione equivale a $\mathfrak{P}$ più lo schema di induzione fino a ε_0. L'idea di considerare successioni di teorie si basa sul fatto seguente: dato un sistema formale $\mathfrak{S}$ possiamo considerarlo nella sua globalità e pensare di rafforzarlo mediante l'introduzione di ulteriori assiomi, in particolare proprio del principio di riflessione; sorge allora naturale l'idea di iterare questo processo nel transfinito recuperando così «pezzo per pezzo» i limiti intrinseci che un singolo sistema o un insieme finito di sistemi presenta a causa dei teoremi di Gödel del 1931. Se riferita allora a questi teoremi e all'aritmetica di Peano la situazione è *grosso modo* la seguente: per recuperare la possibilità di dimostrare la coerenza è necessario – abbiamo visto – l'iterazione transfinita fino a ε_0 di un'operazione effettiva (che può essere identificata con il procedimento di «misurazione» delle dimostrazioni); per recuperare la completezza sintattica il cammino è analogo, ché occorre considerare successioni transfinite di teorie, come sopra detto, muovendosi lungo un segmento della struttura degli ordinali determinato da un ordinale costruttivo (o ricorsivo): si può allora far vedere che $\mathfrak{P}$ è completa in una successione così generata. Come scrive la Dalla Chiara: «In termini un po' imprecisi potremmo riassumere la situazione nel modo seguente: i problemi dell'aritmetica intuitiva non sono tutti risolubili in un unico sistema formale, e nemmeno in una "combinazione" finita di sistemi formali. Tuttavia combinando opportunamente infiniti sistemi formali in una maniera "abbastanza costruttiva" risulta possibile risolvere qualunque problema dell'aritmetica intuitiva».

In tutti e due i casi l'analisi offerta dalla teoria della dimostrazione permetteva, come si verificò ben presto, di approssimare i risultati voluti (coerenza, completezza) misurando i passi necessari.

Non può sfuggire al lettore la connessione di queste ricerche con tutti i tentativi che, sin dal primo svilupparsi della teoria degli ordinali, i costruttivisti (intuizionisti e predicativisti) erano andati conducendo per una analisi della seconda classe cantoriana, con l'obiettivo di isolarne segmenti sempre più ampi dominabili dal punto di vista effettivo. Come già si è ricordato, nel contesto della teoria della dimostrazione queste indagini trovavano una nuova ragion d'essere: da una parte ampi lavori di Feferman sviluppavano la teoria delle *notazioni* di ordinali e dei buoni ordinamenti dimostrabili in modo costruttivo; dall'altra le analisi «informali» di Kreisel ponevano in luce il significato della costruttività degli ordinali e dell'induzione sugli stessi in termini di *visualizzabilità*, giungendo alla conclusione che – almeno in questo senso – il procedimento di induzione fino a ε_0 costituisce il limite delle strutture ad albero che possiamo visualizzare. Con questa identificazione di «visualizzabile» (si pensi agli alberi) con «finitista», Kreisel, pur prendendo come abbiamo visto lo spunto da esigenze «hilbertiane», capovolgeva completamente la direzione dell'approccio hilbertiano stesso, mettendo in luce quella che ne era l'essenziale deficienza: la mancanza di un'analisi adeguata del «concreto», dell'«intuitivo», del «finito». In questo modo risulta chiaro come i metodi elaborati dalla teoria della dimostrazione non solo permettano un'analisi di teorie già formalizzate, ma siano strumenti utili per la ricerca di formalizzazioni adeguate.

Un esempio illuminante al riguardo è costituito dalle ricerche legate alla matematica predicativista e in particolare all'analisi predicativista della dimostrabilità e definibilità. Come abbiamo già avuto occasione di ricordare, la teoria ramificata dei tipi quale si presentava nei *Principia mathematica* una volta eliminato l'assioma di riducibilità, costituiva il punto di riferimento obbligato per ogni tentativo in questo senso. Sul piano linguistico si assumono variabili per ogni tipo finito, cioè variabili per il tipo 0 dei naturali e, posto che $t_1,...,t_n$ siano tipi, variabili per il tipo $(t_1,...,t_n)$ delle relazioni il cui i-esimo argomento è di tipo t_i. Assunta la possibilità di quantificare su variabili di ogni tipo, si possono distinguere le formule in predicative e non, vale a dire – parlando intuitivamente – a seconda che compaiano o meno quantificate variabili del tipo cui appartiene la relazione denotata dalla formula. L'idea di fondo di una concezione predicativista è che mentre i naturali so-

no accettati come dati, gli oggetti di tipo superiore non possono mai essere considerati come una totalità data in atto, ma solo come una totalità potenziale che viene realizzata a livelli. È questo il modello intuitivo riflesso nel linguaggio della teoria ramificata, ma è chiaro anche che limitarsi a procedere semplicemente sostituendo ai tipi semplici il linguaggio ramificato non può che portare a una sorta di parodia della matematica classica, come è stato detto. Il problema, come risultò chiaro almeno a partire dalla metà degli anni cinquanta, era allora quello di trovare formulazioni del predicativismo più adeguate alle concezioni di Poincaré e di Weyl, che fossero al contempo più duttili sia dal punto di vista matematico che da quello metamatematico. Fu questo che portò ricercatori come Lorenzen, Schütte, Wang, Kleene, Spector, ecc. a provare strade diverse che culminarono nei primi anni sessanta in due tipi di proposte, rispettivamente dovute a Kreisel e Feferman.

Il modello che fa da sfondo alle due proposte poggia da una parte su una caratterizzazione della definibilità predicativa, dall'altra su un'esplicitazione, in termini di gerarchie, del concetto di costruzione per livelli. Entrambe le concezioni assumono che una condizione $\mathscr{A}(x)$ ha un significato determinato dal punto di vista predicativista solo quando questo significato è indipendente dall'esistenza o meno dell'insieme degli oggetti che soddisfano la condizione, o in altri termini – riprendendo l'idea di Poincaré – quando il denotato della condizione non cambia una volta che si allarghi l'universo entro il quale lo consideriamo introducendo nuovi oggetti. È in questo modo che emerge un collegamento con gli insiemi iperaritmetici. Se infatti partiamo prendendo come livello di base M_0 la totalità degli insiemi di naturali definibili aritmeticamente (definibili cioè da formule del linguaggio della teoria dei numeri) e definiamo una gerarchia M_α ponendo che

$$M_{\alpha+1} = \{S \subseteq M \mid S \text{ definito predicativamente a partire da } M_\alpha\}$$

e ponendo al solito che M_λ per λ limite è la somma dei passi precedenti, abbiamo un possibile modello degli insiemi di numeri naturali accettabili predicativamente. Il problema centrale era quello di determinare il segmento di ordinali a cui fermarsi, in quanto, come sappiamo, non tutti gli ordinali si possono considerare come

dati in una prospettiva predicativista e senza questo limite la gerarchia coinciderebbe con quella dei costruibili di Gödel. La soluzione sta nel confinarsi a quegli ordinali limite che corrispondono a buoni ordini appartenenti a livelli precedenti nella gerarchia. Questa *condizione di autonomia* è di fondamentale importanza. Nel 1955 Clifford Spector provava che di fatto era sufficiente limitarsi a buoni ordini ricorsivi, quindi a tutti gli ordinali minori del primo ordinale non ricorsivo, indicato con ω_1. Come Kleene avrebbe dimostrato nel 1959, gli insiemi di M_{ω_1} coincidono con gli iperaritmetici, e quindi con quegli insiemi analitici definibili tanto in forma Σ^1_1, quanto Π^1_1. Non stupisce quindi che da più parti si sia affermata l'identità fra gli insiemi iperaritmetici e gli insiemi definibili predicativamente, anche se questo ha dato luogo a discussioni.

La teoria della dimostrazione interviene una volta che si passi al piano della formalizzazione e si cerchi di determinare a livello di teorie ciò che si può definire e dimostrare sulla base del modello precedente. Si giunse così nei primi anni sessanta allo studio delle *progressioni autonome di teorie*, e più precisamente dei due tipi proposti rispettivamente da Feferman e Kreisel. In entrambi i casi abbiamo a che fare con successioni di teorie, rispettivamente R_α e H_α, la prima delle quali incorpora un'analisi ramificata transfinita (essenzialmente una teoria transfinita dei tipi in cui il tipo 0 è quello dei naturali), la seconda basata invece sulla cosiddetta regola HCR (regola di comprensione iperaritmetica) che afferma che se si è provato che una condizione si esprime tanto in forma Σ^1_1 quanto in forma Π^1_1, si può concludere che esiste l'insieme S degli oggetti che la soddisfano, e su una versione formalizzata della ω-regola:

$$\forall x \mathrm{Dim}_{\mathfrak{T}}(\ulcorner \mathscr{A}(x) \urcorner) \rightarrow \forall x \mathscr{A}(x).$$

La progressione in esame viene costruita essenzialmente prendendo come H_0 l'aritmetica al secondo ordine più lo schema di comprensione iperaritmetico e ponendo che $H_{\alpha+1}$ si ottiene da H_α aggiungendo la ω-regola corrispondente a $\mathfrak{T} = H_\alpha$. Per questa progressione come per quella ramificata R_α le condizioni di autonomia vengono formulate in termini di dimostrabilità: alla successione appartiene una teoria H_β se esiste una H_α nella successione che dimostra il principio di induzione $I(z)$ dove z è la notazione per l'ordinale β nel senso di Church-Kleene.

L'idea di considerare progressioni di questo tipo era stata formulata per la prima volta da Turing nel 1939 nel contesto di quelle che egli chiamava *logiche ordinali*. Non stupisce che da indagini di questo genere abbiano preso le mosse tanto ricerche sulle notazioni per ordinali quanto studi dedicati a sottosistemi dell'Analisi basati su schemi di comprensione più deboli di quello classico (quale ad esempio HCR) quanto ancora indagini su regole infinitarie come l'ω-regola. Il fatto importante nel nostro contesto è che le due progressioni di cui sopra abbiamo parlato si dimostrarono sostanzialmente equivalenti dal punto di vista della capacità dimostrativa, offrendo quindi la possibilità di tradurre l'Analisi ramificata all'interno delle più semplici e maneggevoli teorie H_α. Altro risultato interessante al riguardo, legato al programma riduzionista di Hilbert, è la determinazione, ottenuta nei primi anni sessanta da Kreisel e William Tait, del minimo ordinale Γ_0 impredicativo rispetto alla dimostrabilità e l'analisi del suo ruolo nella dimostrazione della coerenza delle teorie determinate dalle successioni di sopra.

Come ricordato sopra però la teoria della dimostrazione non si limitava più al momento riduttivo delle teorie, ma a partire dalla sistemazione di Gentzen si orientava verso un'analisi diretta delle dimostrazioni intuitive. Sullo sfondo – e anche in questo caso decisivo è stato il contributo di Kreisel almeno sul piano analitico e programmatico – rimaneva una domanda più generale che, come appunto osservava Kreisel nel 1970, apriva una prospettiva che per molto tempo era stata negletta. È possibile uno studio sistematico delle forme possibili di dimostrazioni, una loro classificazione «geometrica» che si basi sulla loro struttura ad albero e che sia realmente rilevante ai fini dello studio delle dimostrazioni quali si presentano in matematica?

A partire almeno dal 1965, Dag Prawitz ha sviluppato tutto un programma volto alla costruzione di una *teoria generale della dimostrazione* contrapposta a quella riduttiva di tipo hilbertiano, con lo scopo appunto di affrontare senza preclusioni per quanto riguarda i metodi le dimostrazioni intuitive (opportunamente codificate entro calcoli logici) classificarle, studiarne le possibili relazioni di equivalenza e mutua trasformabilità, indagando quando e perché una dimostrazione è più diretta di un'altra, è più informativa, ecc. In questo modo Prawitz apriva nuovi orizzonti alla teoria della di-

mostrazione affrontando la questione sulla base di una nuova formulazione dei calcoli di Gentzen che egli riuscirà a estendere a un gran numero di diversi sistemi logici (classici e non classici, finitari e non).[5]

Il sistema di Prawitz non è basato sul concetto di sequenza, ma mediante opportune modificazioni riprende l'idea dei calcoli **N** riacquistandone così la «naturalezza» pur non rinunciando ai risultati che Gentzen aveva ottenuto con i più complessi calcoli **L**, in particolare allo *Hauptsatz*, che anzi in questo contesto acquista veramente il significato di *normalizzazione* delle dimostrazioni, vale a dire permette la riconduzione di ogni dimostrazione a una forma diretta, priva di «giri viziosi». Di fatto, l'introduzione dei calcoli **L** da parte di Gentzen era stata motivata proprio dall'impossibilità di formulare in modo esatto un risultato come l'*Hauptsatz* prendendo come codificazione delle dimostrazioni quella dei calcoli **N**, che, pure, erano intuitivamente più motivati. Prawitz riusciva nel 1965 (con la pubblicazione del suo libro *Natural Deduction* [*Deduzione naturale*]) a superare l'ostacolo, sostanzialmente spostando l'attenzione dalla logica classica e da quella intuizionistica, a quella minimale, isolata da Johansson nel 1936.

Il calcolo base per Prawitz è dato da un sistema **M**, di logica minimale appunto, formulato in un linguaggio con &, $\vee$, $\supset$, $\exists$, $\forall$, e una costante proposizionale $\curlywedge$, l'assurdo, mediante la quale è possibile definire la negazione ponendo

$$\neg \mathscr{A} = \mathscr{A} \supset \curlywedge$$

Il calcolo **I** per la logica intuizionista si ottiene da **M** imponendo una condizione sull'assurdo che equivale al principio di Duns Scoto, mentre il calcolo per la logica classica **C** oltre ad imporre l'equivalente del terzo escluso sull'assurdo, richiede una riduzione degli operatori logici assunti come primitivi. Le regole per il calcolo **M** hanno una struttura perfettamente simmetrica: per ogni ope-

[5] Tra i sistemi più interessanti cui Prawitz ha potuto estendere il suo metodo, ricordiamo la logica classica del secondo ordine; la dimostrazione, da lui data, dell'*Hauptsatz* per questo sistema conferma con il suo carattere non costruttivo quanto detto sopra sull'approccio della teoria generale della dimostrazione e sulla sua mancanza di preclusioni verso metodi non elementari, come vedremo più avanti.

ratore logico avremo regole di introduzione e di eliminazione. Esse sono:

$$\&I)\ \frac{\mathscr{A}\ \mathscr{B}}{\mathscr{A}\&\mathscr{B}} \qquad \&E)\ \frac{\mathscr{A}\&\mathscr{B}}{\mathscr{A}} \quad \frac{\mathscr{A}\&\mathscr{B}}{\mathscr{B}}$$

$$\vee I)\ \frac{\mathscr{A}}{\mathscr{A}\vee\mathscr{B}} \quad \frac{\mathscr{B}}{\mathscr{A}\vee\mathscr{B}} \qquad \vee E)\ \frac{\mathscr{A}\vee\mathscr{B}\quad \begin{matrix}[\mathscr{A}]\\ \mathscr{C}\end{matrix} \quad \begin{matrix}[\mathscr{B}]\\ \mathscr{C}\end{matrix}}{\mathscr{C}}$$

$$\supset I\ \frac{\begin{matrix}[\mathscr{A}]\\ \mathscr{B}\end{matrix}}{\mathscr{A}\supset\mathscr{B}} \qquad \supset \mathrm{E})\ \frac{\mathscr{A}\ \mathscr{A}\supset\mathscr{B}}{\mathscr{B}}$$

$$\forall I)\ \frac{\mathscr{A}(a)}{\forall x\mathscr{A}(x)} \qquad \forall E)\ \frac{\forall x\mathscr{A}(x)}{\mathscr{A}(t)}$$

$$\exists I)\ \frac{\mathscr{A}(t)}{\exists x\mathscr{A}(x)} \qquad \exists E)\ \frac{\exists x\mathscr{A}(x)\quad \begin{matrix}[\mathscr{A}(a)]\\ \mathscr{B}\end{matrix}}{\mathscr{B}}$$

dove per le regole $\forall I$ e $\exists E$ occorre osservare le solite restrizioni sulle variabili libere. Per la negazione è facile ricavare, in forza della definizione adottata le regole di introduzione ed eliminazione seguenti

$$\neg\ I)\ \frac{\begin{matrix}[\mathscr{A}]\\ \curlywedge\end{matrix}}{\neg\ \mathscr{A}} \qquad \neg\ E)\ \frac{\mathscr{A}\ \neg\ \mathscr{A}}{\curlywedge}$$

che codificano il fatto che nel calcolo **M** nulla di specifico viene assunto sulle proprietà dell'assurdo. Come si vede, le regole, in alcuni casi, ad esempio in $\supset I$ comportano la messa fra parentesi quadre di alcune formule che occorrono precedentemente nella dimostrazione. Intuitivamente questo significa che le formule in questione vengono *scaricate*, nel senso che la formula dimostrata non dipende da esse. Una dimostrazione in **M** assume naturalmente una disposizione ad albero rovesciato (con l'origine in basso, o, come si usa dire in inglese *upward branching*). Ogni nodo dell'albero è occupato dall'occorrenza di una formula e ogni passo da nodi a

nodi immediatamente inferiori è giustificato dall'applicazione di una regola che si può indicare a fianco. Una formula $\mathscr{A}$ è *dimostrata* nel calcolo minimale **M** dalle assunzioni $\mathscr{A}_1, \ldots, \mathscr{A}_n$ e si scrive

$$\mathscr{A}_1, \ldots, \mathscr{A}_n \vdash_{\mathbf{M}} \mathscr{A}$$

quando esiste un albero come sopra la cui radice è $\mathscr{A}$ e in cui le assunzioni $\mathscr{A}_1, \ldots, \mathscr{A}_n$ sono tutte le formule che compaiono nell'albero, che non sono state scaricate applicando le regole e che non hanno formule che le precedono nell'albero. Due osservazioni sono importanti. La prima è che la rappresentazione delle dimostrazioni come alberi, quindi come oggetti matematici ben definiti, permette di studiare fatti riguardanti le dimostrazioni in modo geometrico, utilizzando concetti quali quello di ramo, diramazione, nodo terminale ecc. La seconda è che nel calcolo **M** le regole o ci permettono di introdurre singoli operatori logici, come per esempio

$$\&I \quad \frac{\mathscr{A} \quad \mathscr{B}}{\mathscr{A} \,\&\, \mathscr{B}}$$

oppure ci consentono di eliminare un operatore logico

$$\&E \quad \frac{\mathscr{A} \quad \mathscr{B}}{\mathscr{A}} \qquad \frac{\mathscr{A} \,\&\, \mathscr{B}}{\mathscr{B}}.$$

Questo è dovuto al fatto che non si è postulato nulla sulla costante $\wedge$.

La situazione cambia quando estendiamo **M** al calcolo **I** per la logica intuizionista aggiungendo la regola

$$\frac{\wedge}{\mathscr{A}}$$

dove $\mathscr{A}$ è *atomica*, oppure al calcolo **C** per la logica classica che si ottiene eliminando le regole per $\wedge$ ed $\exists$ del calcolo **I** e aggiungendo l'ulteriore regola

$$\frac{\begin{matrix}[\neg\mathscr{A}]\\ \vdots\\ \curlywedge\end{matrix}}{\mathscr{A}}$$

dove sempre $\mathscr{A}$ è atomica (e diversa da $\curlywedge$).

Si noti che l'eliminazione di $\vee$ ed $\exists$ non è limitativa in quanto, come sappiamo, dal punto di vista classico possiamo definire entrambi gli operatori ponendo rispettivamente

$$\mathscr{A} \vee \mathscr{B} =_{Df} \neg (\neg\mathscr{A} \wedge \neg \mathscr{B}) \qquad \exists x\mathscr{A} =_{Df} \neg \forall x \neg \mathscr{A}.$$

Le regole su $\curlywedge$ rompono la simmetria tra regole di introduzione ed eliminazione ed il fatto essenziale è che esse vengono assunte *solo* su formule atomiche. In questo modo – diciamo così – esse non vengono a «turbare», nelle singole dimostrazioni, l'intergioco tra introduzione ed eliminazione, in quanto vengono applicate a formule che non contengono operatori logici. Questa tecnica applicata a **M**, **I**, **C** si può estendere, come Prawitz ha mostrato, a sistemi più generali, aggiungendo alle regole del calcolo logico che ci interessa, *sistemi atomici*, ossia regole che riguardano solo formule prive di operatori logici. Esempi significativi sono le regole per l'identità o quelle dei sistemi di Post, che hanno la forma

$$\frac{\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_n}{\mathscr{B}}$$

dove tutte le formule sono atomiche.

Quello che conta è che mediante queste formulazioni delle regole logiche ci è possibile spiegare in che senso una dimostrazione possa non contenere giri viziosi. Abbiamo visto come in un senso le regole di eliminazione siano duali di quelle di introduzione. Questo suggerisce l'idea che mentre le regole di introduzione fissano il significato degli operatori logici, quelle di eliminazione stabiliscono soltanto che cosa deve valere in base al *solo* significato della premessa maggiore (vale a dire quella che contiene l'operatore logico eliminato). Come scrive Prawitz «in altre parole, una dimostrazione della conclusione di una eliminazione è già "contenuta" nelle dimostrazioni delle premesse quando la premessa maggiore è derivata mediante introduzione». È questo quel *princi-*

pio di inversione fra regole di eliminazione e di introduzione che Gentzen aveva certamente intravisto ma che non era in grado di formulare *direttamente* sui calcoli **N**. Si trattava di una inadeguatezza delle dimostrazioni formali rispetto a quelle intuitive che i sistemi di Prawitz eliminano mostrando appunto come sia possibile un'analisi diretta delle *dimostrazioni* e non solo della *dimostrabilità*, una volta che si utilizzino gli strumenti formali adeguati. La prova di questo fatto sta nella possibilità di precisare la portata del principio di inversione mostrando in modo esatto come ogni dimostrazione sia trasformabile mediante precise operazioni sintattiche in un'altra che non contiene giri viziosi nel senso specificato dal principio di inversione, vale a dire che non elimina un operatore sfruttando esattamente quanto ci ha permesso di introdurlo precedentemente in una dimostrazione.

La forma tipica di una dimostrazione in cui non occorrono giri viziosi è data da un albero in cui, partendo dall'alto, si decompongono le assunzioni applicando solo regole di eliminazione (parte *analitica* della dimostrazione) si applicano eventualmente regole *atomiche* come quella per $\curlywedge$ alle formule indecomponibili ottenute (parte *minima* della dimostrazione) ed infine si ottengono formule composte utilizzando le regole di introduzione (parte *sintetica* della dimostrazione).

Chiamiamo *normali* le dimostrazioni che hanno questa forma. Il teorema di normalizzazione, che Prawitz dimostra per **M**, **I** e **C** afferma allora che *ogni* dimostrazione può essere trasformata in una dimostrazione normale; che questo non sia affatto banale lo si vede subito osservando che l'eliminazione di un giro vizioso può comportare l'introduzione di un altro e che quindi la dimostrazione che ogni giro vizioso può essere eliminato avviene attraverso un delicato processo di induzione in cui risulta estremamente vantaggioso il fatto che le dimostrazioni abbiano struttura ad albero (il che permette l'assegnazione di gradi ordinali ai nodi e all'applicazione di regole). La differenza fra un *teorema di forma normale* e un *teorema di normalizzazione* sta nel fatto che il primo ha un puro carattere esistenziale («esiste una dimostrazione normale») mentre il secondo specifica che la dimostrazione in questione si ottiene a partire da quella data mediante prefissate regole di trasformazione.

Il processo di trasformazione di una dimostrazione Σ in una Σ'

in forma normale avviene applicando – come abbiamo detto – specifiche operazioni del tipo ad esempio

$$\dfrac{\dfrac{\Sigma_1}{\mathscr{A}_1} \quad \dfrac{\begin{matrix}[\mathscr{A}_1] \\ \Sigma_2 \\ \hline \mathscr{A}_2\end{matrix}}{\mathscr{A}_1 \supset \mathscr{A}_2}}{\mathscr{A}_2} \qquad \dfrac{\begin{matrix}\dfrac{\Sigma_1}{[\mathscr{A}_1]} \\ \Sigma_2\end{matrix}}{\mathscr{A}_2}$$

(qui Σ_1, Σ_2 rappresentano dimostrazioni e $[\mathscr{A}]$ denota l'insieme di assunzioni in Σ_2 che sono chiuse dalla $\supset I$) in cui appunto si elimina un'introduzione che sia seguita immediatamente da una eliminazione (nell'esempio per il connettivo implica). Queste operazioni si chiamano *riduzioni* e si affiancano ad altre operazioni che possiamo considerare, le *semplificazioni*, come ad esempio

$$\dfrac{\dfrac{\Sigma}{\mathscr{A}_1 \vee \mathscr{A}_2} \quad \dfrac{\Sigma_1}{\mathscr{B}} \quad \dfrac{\Sigma_2}{\mathscr{B}}}{\mathscr{B}} \qquad \dfrac{\Sigma_i}{\mathscr{B}}$$

in cui si eliminano applicazioni ridondanti di $\vee E$, e le *espansioni* come ad esempio la seguente

$$\dfrac{\Sigma}{\mathscr{A} \,\&\, \mathscr{B}} \qquad \dfrac{\dfrac{\dfrac{\Sigma}{\mathscr{A} \,\&\, \mathscr{B}}}{\mathscr{A}} \quad \dfrac{\dfrac{\Sigma}{\mathscr{A} \,\&\, \mathscr{B}}}{\mathscr{B}}}{\mathscr{A} \,\&\, \mathscr{B}}$$

che ci permettono di avere formule minimali solo atomiche.

Disponiamo così di una sorta di algebra delle dimostrazioni, in cui le dimostrazioni sono oggetto di indagine e le trasformazioni sopra viste ci permettono di manipolarle (si noti: in modo costruttivo) e di collegarle l'un l'altra. Possiamo così definire il concetto di riducibilità fra dimostrazioni e dire che Σ si riduce a Σ' quando esiste una successione $\Sigma_1, \ldots, \Sigma_n$ di dimostrazioni dove Σ_1 è Σ, Σ_n è Σ' e ogni Σ_{i+1} è ottenuta da Σ_i mediante una riduzione. Può capitare che una dimostrazione non sia riducibile ad altre che a se

stessa e in questo caso la chiamiamo *irriducibile*; il fatto importante è che una dimostrazione è irriducibile se e solo se è in forma normale. Il teorema di normalizzazione si può allora formulare dicendo che ogni dimostrazione nei nostri calcoli si riduce ad una dimostrazione irriducibile. Ma c'è di più: possiamo ottenere un teorema *forte* di normalizzazione che ci dice che ogni dimostrazione si riduce ad un'*unica* forma normale e che ogni successione di passi di riduzione che parte da una dimostrazione terminerà con la sua unica forma normale.

In questo modo ogni dimostrazione intuitiva viene ad avere una sorta di forma canonica ed è naturale porsi la domanda di quando due dimostrazioni si possano considerare la *stessa* dimostrazione. È questo un problema che si presenta spesso nella matematica (e oggi nella *computer science*). Uno stesso teorema si può ottenere in modi diversi; a volte queste differenze sono profonde, in quanto coinvolgono principi di forza differente e addirittura di contenuto concettuale diverso. A volte invece le differenze sono puramente «logiche», di stile formale, ed è questa «identità» che si può cercare di indagare utilizzando i risultati di normalizzazione.

Possiamo dire *equivalenti* due dimostrazioni Σ e Σ' quando si ottengono l'una dall'altra mediante riduzione (o l'inverso di una riduzione) e l'idea è che dimostrazioni simili differiscono solo per l'eventuale presenza nell'una di qualche giro vizioso che non figura nell'altra. La congettura che Prawitz formulò nel 1971 è che di fatto due dimostrazioni siano la stessa dimostrazione se e solo se sono equivalenti nel senso sopra detto e una certa plausibilità è conferita a questa ipotesi dal teorema di normalizzazione forte che ci dice in questo caso che due dimostrazioni normali differenti non sono mai equivalenti.

La portata della teoria della dimostrazione non riduttiva nel senso di Prawitz non si limita a questo tipo di risultati di carattere più filosofico. Con questi metodi Prawitz è riuscito a dimostrare in modo unitario tutti i risultati sulle traduzioni tra logica minimale, intuizionistica e classica di cui abbiamo parlato a suo tempo aggiungendovene di nuovi e ha dimostrato come corollario del teorema di normalizzazione siano le *proprietà della disgiunzione* e *dell'esistenziale* per **I** che affermano rispettivamente che se $\mathscr{A} \vee \mathscr{B}$ è dimostrabile allora o lo è $\mathscr{A}$ o lo è $\mathscr{B}$ e che se è dimostrabile $\exists x\, \mathscr{A}$ allora è dimostrabile anche $\mathscr{A}(t)$ per un certo termine t. Proprietà di que-

sto tipo sono estremamente interessanti, ove si passi dai calcoli puri a teorie matematicamente più ricche, perché ci permettono di ottenere informazioni di tipo costruttivo; ad esempio nel caso della proprietà dell'esistenziale, ci garantiscono termini specifici che realizzano la formula. Ciò spiega come sia stata molto attiva in tempi recenti la ricerca di teorie che godano di queste proprietà. Classicamente solo le teorie complete ne godono, ma intuizionisticamente, come ha mostrato Harvey Friedman, questo succede non solo per la logica pura ma per ampie classi di estensioni dell'aritmetica intuizionista ed in specifici contesti la proprietà della disgiunzione implica quella dell'esistenziale.

La concezione della teoria della dimostrazione introdotta da Prawitz ha portato soprattutto negli anni sessanta a una notevole messe di risultati nuovi e interessanti che si incentrano tutti sui risultati di normalizzazione. Teoremi di questo tipo valgono ad esempio per l'aritmetica di Peano al primo ordine (qualunque logica delle tre sopra nominate si utilizzi) una volta che si impieghi come regola di deduzione la seguente *regola di induzione*

$$\frac{A(0) \quad \begin{matrix}[A(a)] \\ A(a')\end{matrix}}{A(t)}$$

dove a' indica il successore di a, come è stato dimostrato da H. Jervell nel 1970. Al 1969 invece risalgono alcune ricerche di P. Martin-Löf sulla logica proposizionale infinitaria, minimale o intuizionista per cui viene stabilito un teorema di normalizzazione, cosa che Martin-Löf farà negli anni successivi anche per la teoria delle definizioni induttive iterate. Va ricordato che naturalmente a livello metateoretico quando si passa a provare i teoremi di normalizzazione per sistemi più forti della logica pura le procedure induttive utilizzate coinvolgono elementi che non possiamo assimilare alla matematica finitista, come succedeva per la induzione fino a ε_0 utilizzata da Gentzen. Si tratta di un'ulteriore conferma del mutato atteggiamento rispetto ad Hilbert: il ricorso a principi non finitisti non avviene *faute de mieux*, ma si basa sulla convinzione che le dimostrazioni (e non la semplice dimostrabilità) sono oggetti di per sé interessanti, a prescindere da qualunque contesto di tipo riduttivo o meno.

Quanto detto non va inteso ovviamente nel senso che tanto la teoria riduzionista della dimostrazione che in generale le ricerche iniziate da Gentzen abbiano perduto interesse. Il nodo metodi metateorici/*Hauptsatz* o teoremi di normalizzazione per sistemi forti ha un interesse intrinseco fondamentale in quanto, qualunque sia l'atteggiamento fondazionale che si assume – sia che ci si interessi di metodi pratici per ottenere informazioni di tipo computazionale, sia che si indaghi l'effettivo peso dei principi infinitari nella pratica matematica – ad un'analisi della struttura delle dimostrazioni di una teoria non si può sfuggire. Al 1974 risale un'indagine ravvicinata di J. Zucker che confronta eliminazione del taglio e normalizzazione stabilendo un ponte fra i due approcci, cosicché i problemi formulati in termini dell'uno si possono ritradurre in termini dell'altro.

Un tipico esempio di problema formulato nel contesto della teoria della dimostrazione alla Gentzen è la famosa *congettura di Takeuti*, del 1953, che afferma l'esistenza di una forma senza tagli per ogni dimostrazione nel calcolo delle sequenze per la logica classica di ordine ω, vale a dire la teoria dei tipi semplici. La congettura fu dimostrata negli anni 1966-68 da W. Tait, M. Takahashi e D. Prawitz con metodi semantici e quindi in maniera decisamente non costruttiva partendo da una riformulazione semantica data da K. Schütte nel 1960 sulla base di un interessantissimo metodo in termini di valutazioni parziali.

Con metodi analoghi nel 1968 Prawitz e nel 1970 Takahashi estesero il risultato al caso intuizionista della teoria della specie, l'analogo della classica teoria degli insiemi. L'importanza di questi risultati non va sottovalutata in quanto sostanzialmente questi sistemi (con eventuali aggiunte di regole atomiche) costituiscono il contesto naturale in cui ricostruire l'Analisi. Si spiegano quindi i tentativi di perfezionarli mediante riformulazioni di carattere più costruttivo. Così ad esempio Takeuti nel 1967, utilizzando la sua tecnica dei diagrammi ordinali, riuscì a dimostrare l'*Hauptsatz* per una sottoteoria significativa della logica classica del secondo ordine in cui il principio di comprensione viene ristretto a formule del tipo Π^1_1 impiegando solo, a livello metateorico, il principio di induzione su certi ordinali ricorsivi. Da questo risultato scendeva la coerenza dell'Analisi, quindi un risultato significativo anche se non accettabile nel contesto della teoria riduttiva della dimostrazione.

Partendo, invece che dall'analisi delle dimostrazioni di Gentzen, dalla teoria dei funzionali ricorsivi di tipo finito utilizzata da Gödel nel 1968 per analizzare l'aritmetica, Clifford Spector (1930-61) era già riuscito, in un lavoro fondamentale pubblicato postumo nel 1962, a provare la coerenza relativa dell'intera Analisi classica utilizzando principi non finitisti, ma in qualche modo costruttivi. L'idea di Spector era di interpretare l'Analisi classica in una teoria Σ dei funzionali di tipo finito, che estendeva la teoria T di Gödel, introducendo un nuovo schema definitorio, quello di *ricorsione sbarrata*, e assumendo il corrispondente principio di induzione. Questo principio non è finitista ma era stato introdotto come si ricorderà da Brouwer nei suoi studi sulle successioni di libera scelta e costituisce quindi in qualche modo un'assunzione di carattere costruttivo.

Vedremo più avanti come gli anni settanta si aprano con risultati importanti su questi temi che da una parte porteranno a teoremi non solo di forma normale ma di normalizzazione per i sistemi citati, dall'altra, proseguendo lungo la linea aperta da Spector, daranno significato più incisivo all'impiego dei funzionali nella teoria della dimostrazione.

4. LA METAMATEMATICA DELL'INTUIZIONISMO

Abbiamo già dato a suo tempo una breve descrizione dei concetti fondamentali della matematica intuizionista quali si erano andati sviluppando dopo i primi interventi di Brouwer. Se sulle prime il confronto con la matematica classica era risultato spesso difficile e oscurato da pregiudizi ideologici, dopo gli interventi di Heyting negli anni trenta non stupisce che le indagini in questa direzione si siano moltiplicate, gettando luce da una parte sulla forza dei sistemi intuizionisti e sulle possibili alternative che le proposte di Brouwer ammettevano, dall'altra sui rapporti con la matematica classica. Ci sembra difficile esprimere la cosa più incisivamente e chiaramente di quanto non facciano le seguenti parole di John Myhill (1968): «In generale la ripulsa contro la formalizzazione è stata di recente eliminata tra gli intuizionisti anche in Olanda [patria di Brouwer] e i risultati ottenuti impiegandola hanno contribuito grandemente alla chiarificazione delle oscurità.

Ovviamente questa chiarificazione è un compito in prima istanza *filosofico* piuttosto che matematico; ma la formalizzazione vi contribuisce in modo essenziale perché risulta che sostanzialmente l'intero *corpus* brouweriano è derivabile in modo standard (ossia nel calcolo dei predicati intuizionista) da un numero limitato di enunciati... esplicitamente accettati da Brouwer. Il compito filosofico viene così ridotto per mezzo della formalizzazione da quello estremamente impegnativo di "dare senso a" tutti gli scritti di Brouwer a quello più trattabile di "dare senso a" questo piccolo numero di principi ammesso esplicitamente assieme con i principi dello stesso calcolo dei predicati intuizionista; i due tipi di principi verrebbero quindi assunti come assiomi di un sistema formale deduttivo e il resto del lavoro consisterebbe semplicemente di deduzioni da questi assiomi. Ciò non vuol significare che esista finora un accordo generale su quelli che dovrebbero essere gli assiomi (non logici) e questo per due ragioni: (α) se si fa una lista di assiomi sufficienti per la derivazione del *corpus* brouweriano, si trovano varie questioni *semplicissime* per le quali si vorrebbe conoscere la risposta (che in effetti dovrebbe essere conosciuta per comprendere chiaramente i presupposti filosofici di Brouwer) e sulle quali egli essenzialmente non si era impegnato e (β) si hanno prove che le sue concezioni sono cambiate radicalmente in suoi scritti posteriori... o almeno che egli usava le stesse parole (ad esempio "successione di libera scelta") con significati differenti sicché la questione se *tutto* il suo lavoro possa essere compreso in un sistema formale (o, più fondamentalmente, in un unico schema filosofico) è discutibile e difficile e la risposta, sulla base delle indicazioni di cui a tutt'oggi disponiamo, sembra essere affermativa ma decisamente non banale. (Non voglio qui intendere che vi sia coinvolta una matematica difficile, bensì semplicemente un pensiero difficile.)»

Non avrebbe senso dal nostro punto di vista fermarci qui ad analizzare in dettaglio le specifiche assiomatizzazioni che sono state date, a partire dal secondo dopoguerra, dei vari rami della matematica intuizionista. Quello che ci preme sottolineare è che la possibilità di usare metodi metateorici nell'analisi di queste teorie ha avuto un duplice significato: da una parte ha permesso di mettere ordine all'interno della tradizione intuizionista stessa, saggiando le capacità di diverse formulazioni ed indagando la forza di diversi principi (si pensi, per far solo un esempio, ai diversi

principi di continuità compatibili con le concezioni di Brouwer e all'intricata situazione della teoria del soggetto creativo); dall'altra ha reso possibile l'emergere di alternative in ambito costruttivista mostrando come, in diversi contesti, altri modi diversi da quelli di Brouwer si possano dare per sviluppare una matematica costruttiva.

Forse la più antica di queste alternative è costituita dal predicativismo, che dopo gli interventi di Poincaré aveva trovato nel 1918 il suo manifesto nell'opera di H. Weyl. Anche se nell'intuizionismo stesso e in altre correnti costruttiviste possono comparire restrizioni di tipo predicativista per quanto riguarda la definizione di enti matematici, ciò che caratterizza il predicativismo in sé e per sé come corrente rispetto all'intuizionismo è la sua accettazione della logica classica. L'atteggiamento di fondo è quello di costruire tutta intera la matematica a partire dall'insieme dei naturali, la cui totalità viene accettata, allargando man mano il dominio mediante l'uso di definizioni predicative senza cioè ricorrere alla quantificazione riferentesi a totalità che non siano già state costruite. L'ostacolo maggiore è dato dalla presenza, già nell'Analisi elementare, di concetti, tipo quello di confine superiore di un insieme, che non sono predicativi ma che si possono rendere tali ricorrendo a costruzioni induttive definite su segmenti di ordinali.

Proprio sulla possibilità di utilizzare metodi costruttivi, come le definizioni induttive generalizzate nella ricostruzione dell'Analisi, si ebbero negli anni cinquanta interventi come quelli di Hao Wang, Paul Lorenzen e John Myhill che differivano per quanto riguardava il significato di questi metodi in una prospettiva predicativista. Ciò che comunque già i lavori di Weyl avevano mostrato era che sorprendentemente l'intera Analisi delle funzioni continue poteva essere ricostruita in termini predicativisti; il problema sorgeva soprattutto quando si trattava di fare appello a principi come quello del minimo confine superiore chiaramente non accettabili – così come sono – per un predicativista. A partire dalla metà degli anni cinquanta l'approccio di Weyl fu ripreso da Grzegorczyck, Kondo e Kreisel, il quale ultimo pose in luce come ulteriori segmenti dell'Analisi reale potessero essere sviluppati su base aritmetica mediante un'opportuna precisazione di quel che significa «dare» una successione, una funzione, ecc., da una prospettiva predicativista.

Un aspetto diverso di questa attenzione per il costruttivo sviluppava Lorenzen con la sua *logica operativa* e il suo metodo dei dialoghi, che permettono una formalizzazione della logica fondata sull'idea che il significato di connettivi e quantificatori si può rendere esplicito codificando le regole che in un dialogo tra due interlocutori ciascuno di essi utilizza per cercare di refutare l'altro. C'è uno stretto legame fra queste tecniche e quelle di Beth e di Hintikka e la differenza centrale sta nel fatto che nell'approccio di Lorenzen la logica che risulta non è quella classica, ma una logica di tipo intuizionista.

Sempre agli anni cinquanta risalgono anche le prime indagini sviluppate da Markov sulla matematica costruttiva ricorsiva che da allora ha avuto un grande sviluppo soprattutto in Unione Sovietica. Per Markov e la sua scuola oggetti matematici possono essere soli gli algoritmi, i quali vengono identificati, in base alla tesi di Church, con le funzioni ricorsive (si ricordi che una delle precisazioni equivalenti del concetto di ricorsivo è quella data da Markov in termini dei suoi algoritmi normali), le procedure date in termini combinatori che definiscono funzioni, relazioni, ecc. Su questi oggetti si ragiona classicamente salvo in presenza di esistenziali e di disgiunzioni, che devono essere realizzati in modo ricorsivo. Il principio distintivo del costruttivismo nel senso di Markov, è dato dall'accettazione della tesi (detta appunto di Markov, di cui abbiamo parlato in V, 4) secondo la quale se non riusciamo a dimostrare che una computazione non termina, allora possiamo assumere che termini. La cosa è ovvia dal punto di vista classico, ma non in ogni contesto viene accettata dal punto di vista intuizionista.

Avremo occasione più avanti di parlare di altre posizioni di tipo costruttivista non assimilabili all'intuizionismo, ad esempio quella di Bishop. Quello su cui vogliamo soffermarci è un'altra conseguenza che la formalizzazione della matematica intuizionista ha comportato: la possibilità di dare interpretazioni non costruttive – appunto nell'ambito della matematica classica – della matematica intuizionista. Ci riferiamo in particolare a tutta quella serie di lavori che a partire dalle ricerche condotte da Kleene nel 1945 hanno portato al concetto di *realizzabilità ricorsiva* come corrispettivo della validità intuizionista. L'idea di fondo di Kleene è che dal punto di vista intuizionista ogni proposizione è una comunicazio-

ne incompleta riguardante l'esistenza di costruzioni e che quindi la verità di una formula del tipo

$$\forall x \exists y \mathscr{A}(x, y) \qquad (*)$$

deve in qualche modo significare che esiste un metodo per cui dato un qualunque oggetto x si trova un oggetto y per cui $\mathscr{A}(x, y)$ è vera.

Fu del tutto naturale per Kleene, che proprio in quegli anni stava gettando le fondamenta per la teoria generale della ricorsività, vedere se era possibile identificare questo metodo con una funzione ricorsiva (almeno nel caso ci si occupi di formule dell'aritmetica) così da ottenere un risultato analogo al teorema di Bernays del 1939 sulle formule del tipo $\forall\exists$ (si veda la nota 27 a pag. 624). Per poter far questo era essenziale non limitarsi alle formule di questo tipo, ma considerare *tutte* le formule, in quanto non c'è limitazione sulla forma delle proposizioni che possono occorrere in una dimostrazione di una formula come la (*). A questo punto interveniva l'interpretazione degli operatori logici intuizionisti data da Kolmogorov e da Heyting e si trattava di rendere esplicito in termini di funzioni ricorsive il significato di «costruzione».

Nel suo articolo del 1945 dal titolo *On the Interpretation of intuitionistic Number Theory* (*Sull'interpretazione della teoria dei numeri intuizionista*) Kleene giungeva così alla definizione seguente in cui identificando le funzioni ricorsive coi loro indici, la relazione «e realizza $\mathscr{A}$» può essere definita tra numeri naturali e formule chiuse:

1) *e realizza* una formula atomica chiusa $\mathscr{A}$ se $e = 0$ e $\mathscr{A}$ è vera (ossia, se $e = 0$ e $\mathscr{A}$).
2) *e realizza* $\mathscr{A} \wedge \mathscr{B}$ se $e = 2^a \cdot 3^b$, dove *a realizza* $\mathscr{A}$ e *b realizza* $\mathscr{B}$.
3) *e realizza* $\mathscr{A} \vee \mathscr{B}$ se $e = 2^0 \cdot 3^a$, dove *a realizza* $\mathscr{A}$, oppure $e = 2^1 \cdot 3^b$ dove *b realizza* $\mathscr{B}$.
4) *e realizza* $\mathscr{A} \rightarrow \mathscr{B}$ se e è il numero di Gödel di una funzione parziale ricorsiva φ in una variabile tale che, se *a realizza* $\mathscr{A}$, allora $\varphi(a)$ *realizza* $\mathscr{B}$.
5) *e realizza* $\neg\, \mathscr{A}$ se *e realizza* $\mathscr{A} \rightarrow 1 = 0$.
6) *e realizza* $\exists x \mathscr{A}(x)$, se $e = 2^x \cdot 3^a$ dove *a realizza* $\mathscr{A}(\overline{x})$.
7) *e realizza* $\forall x \mathscr{A}(x)$, se e è il numero di Gödel di una funzione ricorsiva generale φ in una variabile tale che, per ogni x, $\varphi(x)$ *realizza* $\mathscr{A}(\overline{x})$.

Una formula chiusa $\mathscr{A}$ del linguaggio della teoria dei numeri si dirà allora *realizzabile* se esiste un numero *e* che la realizza. La definizione si estende a formule arbitrarie ricorrendo alla chiusura universale.

Si impongono due osservazioni. La prima è che è essenziale, nella definizione, considerare funzioni non necessariamente totali. Se si vuole infatti dimostrare che ogni teorema dell'aritmetica intuizionista è realizzabile, è necessario, nella clausola su $\rightarrow$, richiedere solo una funzione che applicata ad argomenti *a* che realizzano $\mathscr{A}$, ci dia valori *b* che realizzano $\mathscr{B}$; per il resto la funzione può essere indefinita ed è questo un esempio tipico dell'importanza del concetto di funzione ricorsiva parziale che proprio Kleene aveva introdotto. L'altro fatto da osservare è che le clausole sulla realizzabilità di per sé non forniscono una *definizione* del significato costruttivo degli operatori logici, in quanto le clausole definitorie contengono gli operatori logici che dovrebbero definire. In altre parole, l'uso della realizzabilità è strumentale, per indagare le proprietà metamatematiche delle teorie intuizioniste e non è un'*analisi* in termini primitivi del concetto di verità intuizionista.

Fu appunto in quest'ottica che Kleene utilizzò il suo teorema per cui se $\mathscr{A}$ è dimostrabile nell'aritmetica intuizionista $\mathscr{A}$ è realizzabile. Oltre a fornire ovviamente una dimostrazione di coerenza dell'aritmetica intuizionista (si può infatti provare – non finitisticamente, beninteso – che sia $\mathscr{A}$ che $\neg\, \mathscr{A}$ non sono simultaneamente realizzabili) il teorema permetteva di concludere che alcune formule, come alcuni esempi del principio del minimo, vere classicamente non sono dimostrabili intuizionisticamente, che esistono estensioni coerenti dell'aritmetica intuizionista che non sono coerenti con l'aritmetica classica ed infine che formule predicative del tipo ad esempio

$$\neg\, \neg\, \forall x(\mathscr{A}(x) \vee \neg\, \mathscr{A}(x))$$

non sono dimostrabili nella logica intuizionista. Quest'ultimo fatto poneva naturalmente il problema se, limitandosi alla logica proposizionale, il calcolo intuizionista fosse completo rispetto alla interpretazione in termini di realizzabilità. Nel 1953 G. Rose dimostrò che così non era e che esistevano formule realizzabili non dimostrabili.

A questa definizione di realizzabilità per la logica e l'aritmetica ne seguirono altre, attorno agli anni sessanta, motivate dal desiderio di estendere l'ambito di applicazione del metodo. Partendo da alcuni risultati di Ronald Harrop del 1956, Kleene si chiese se era possibile dimostrare, via realizzabilità, la validità della proprietà della disgiunzione e dell'esistenziale per l'aritmetica intuizionista. Questo portò Kleene a modificare le clausole sull'$\vee$ e sull'$\exists$ e a introdurre, nel 1960, la nozione di $(\vdash)$ – realizzabilità secondo la quale

e $(\vdash)$ – *realizza* $\mathscr{A} \vee \mathscr{B}$ se e solo se $(e)_0 = 0$ e $(e)_1$ $(\vdash)$ – realizza $\mathscr{A}$ e $\vdash \mathscr{A}$, oppure $(e)_0 \neq 0$ e $(e)_1$ $(\vdash)$ – realizza $\mathscr{B}$ e $\vdash \mathscr{B}$;
e $(\vdash)$ – *realizza* $\exists x(\mathscr{A}(x))$ sse per $x = (e)_0$, $(e)_1$ $(\vdash)$ – realizza $\mathscr{A}(\bar{x})$ e $\vdash \mathscr{A}(\bar{x})$.

il che permette di passare dalla realizzabilità alla dimostrabilità e non solo viceversa. Anche in questo caso Kleene ridimostrava che ogni teorema dell'aritmetica intuizionista è $(\vdash)$ – realizzabile e otteneva per altra via i risultati di Harrop.

L'ultima definizione presenta un miscuglio di condizioni che riguardano la dimostrabilità e le funzioni ricorsive. Come già nel 1962 Kleene poneva in luce, questi due aspetti sono separabili ed è possibile definire una nuova nozione, quella di «slash» indicata con $|$, abbandonando le clausole riguardanti le funzioni. Questa nuova nozione permette ancora di ottenere risultati di dimostrabilità e non dimostrabilità come i precedenti in modo estremamente diretto e verrà analizzata nel 1966 da P. Aczel in termini semantici utilizzando quelle strutture di Kripke di cui ci occuperemo tra poco.

Tornando alla realizzabilità, già dal 1950 Kleene tentò di estenderla all'Analisi intuizionista, quindi al secondo ordine, il che comportava la considerazione di quantificatori su funzioni. Come base assiomatica per questa teoria, Kleene prendeva i sistemi formulati nel suo libro del 1965, scritto in collaborazione con R. Vesley e che è stato una pietra miliare nella storia della formalizzazione della matematica intuizionista: *The Foundations of Intuitionistic Mathematics* (*I fondamenti della matematica intuizionista*). Dopo diverse esitazioni, nel 1959 Kleene giungeva ad una definizione di realizzabilità che riguardava non più numeri (indici per funzioni) ma funzioni numeriche ricorsive (che codificano funzionali, la cui teo-

ria in quegli anni, egli andava sviluppando). Con questa ultima nozione di rappresentabilità, Kleene poteva provare la coerenza dell'Analisi intuizionista (non banale, in quanto vi compaiono principi, come quello di Brouwer, incompatibili con la matematica classica) e l'indipendenza del principio di Markov.

Successivamente altri ricercatori, soprattutto Joan Rand Moschovakis nel 1967 e H. Friedman nel 1971, hanno esteso le applicazioni del metodo che si è andato sempre più rivelando come uno strumento essenziale nello studio della matematica intuizionista.

5. LOGICHE NON CLASSICHE

Diversamente che nel caso della *matematica* intuizionista fu ponendosi più decisamente in una prospettiva affine a quella della teoria dei modelli che risultò possibile dare una sistemazione soddisfacente – almeno dal punto di vista classico, beninteso – alla semantica per la logica intuizionista. Questo si deve in particolare a Saul Kripke che nel 1965, nell'ambito di un'analisi generale di un'ampia famiglia di logiche non classiche, quelle modali, risolse il problema prendendo lo spunto dai sistemi di Lewis già precedentemente analizzati nel 1959. Pressoché simultaneamente anche Hintikka e Stig Kanger giungevano a soluzioni analoghe partendo da versioni analitiche (nel senso illustrato parlando delle tavole di Beth) delle *descrizioni di stato* con cui R. Carnap aveva affrontato quantificazione e modalità nel suo libro fondamentale *Meaning and Necessity* (*Significato e necessità*) del 1947. Il legame fra logica modale e intuizionista non deve stupire se si pensa alla traducibilità della logica intuizionista in quella modale (**S**4 di Lewis) posta in luce da Gödel e Tarski negli anni trenta. Dato il ruolo centrale che questa semantica ha nelle ricerche odierne dei sistemi di logiche intensionali, è opportuno soffermarci brevemente sull'idea di fondo della sistemazione di Kripke.

Partendo dalla nozione di intensione introdotta da Carnap nel 1947, Kripke vede negli operatori modali delle funzioni che si applicano non già all'estensione di un enunciato (il suo valore di verità) bensì alla sua intensione; quest'ultima viene riguardata come una funzione che associa, in corrispondenza a diverse circostanze,

a diversi «mondi possibili», diverse estensioni, diversi valori di verità. Discostandosi da Carnap, però, Kripke non concepisce l'insieme dei mondi possibili come determinato una volta per tutte dall'insieme di tutte le descrizioni coerenti di stati possibili, ma considera più astrattamente i mondi possibili come *indici* scelti in un insieme prefissato. In altre parole Kripke separa il *supporto* I di un'interpretazione $\mathfrak{M}$ in cui compaiono enti astratti come gli indici, da una *funzione* di interpretazione M che ad ogni $i \in I$ associa una valutazione classica M_i nel caso proposizionale, o una struttura classica nel caso del primo ordine. Per comodità quando parleremo di verità (classica) di una formula $\mathscr{A}$ a un indice $i \in I$ intenderemo sempre che $\mathscr{A}$ è vera in M_i, cioè $M_i \models \mathscr{A}$; ma il lettore tenga ben presente che su uno stesso supporto si possono definire diverse funzioni di interpretazione per cui sarebbe più corretto parlare di verità a M_i. Posto questo, intuitivamente possiamo considerare come *intensione* di una formula $\mathscr{A}$ in un'interpretazione l'insieme dei livelli in cui $\mathscr{A}$ è vera, o se si vuole la funzione che ad ogni livello assegna valore 0 o 1 a seconda che la formula sia falsa o meno.

Per assegnare all'enunciato $\mathscr{A}$ la sua intensione $f_{\mathscr{A}}$, cioè l'insieme dei mondi in cui esso è vero (vale a dire quegli $i \in I$ per cui $\mathfrak{M} \models_i \mathscr{A}$) basterà procedere induttivamente utilizzando le clausole classiche per $\wedge$, $\vee$, $\neg$, $\rightarrow$ e valutare $\Box$ facendo riferimento all'insieme I dei mondi possibili. Che cosa significherà affermare che l'enunciato $\Box\, \mathscr{A}$ è vero? In prima istanza si può dire che la necessità di $\mathscr{A}$ coincide con la sua verità in ogni mondo possibile. Questa nozione però risulta troppo grossolana e ammette una importante precisazione: i mondi possibili non sono tutti *compossibili*, nel senso che dato un mondo $i \in I$, solo alcuni degli altri elementi di I sono «accessibili» da i, cioè sono possibili dal punto di vista di i. Ciò posto, per assegnare ad ogni enunciato una intensione occorrerà non solo specificare l'insieme I dei mondi possibili, ma anche una relazione binaria R fra di essi che intenderemo come relazione di accessibilità; iRj andrà cioè letto: *il mondo j è accessibile dal mondo i*, o in altre parole j è un mondo possibile per i; chiameremo quindi *struttura modello* ogni coppia $\mathfrak{I} = < I, R >$ e le interpretazioni saranno coppie $\mathfrak{M} = < \mathfrak{I}, M >$. Porremo allora che $\mathfrak{M} \models_i \Box\, \mathscr{A}$ nel caso $\mathfrak{M} \models_j \mathscr{A}$ per ogni $j \in J$ per cui iRj.

Il grande interesse di questa analisi sta nel fatto che permette di dare una sistemazione unitaria ad un'enorme varietà di nozioni intensionali. Imponendo infatti proprietà particolari (riflessività e/o simmetria e/o transitività, ecc.) alla relazione R è possibile stringere più da vicino le diverse idee di necessità che abbiamo intuitivamente. Così se intendiamo la necessità come verità in ogni istante, interpreteremo I come l'insieme degli istanti ed R come la relazione di (pre)ordine (ossia riflessiva e transitiva) che sussiste fra istanti di tempo. Non deve stupire quindi che partendo da questa idea sia possibile ottenere delle semantiche intuitive ed estremamente duttili per i sistemi modali presentati da Lewis e altri, nel senso che alle diverse nozioni di necessità canonizzate nei vari sistemi **S**1-**S**5 corrispondono diverse proprietà della relazione R: ad esempio, al sistema **S**4 corrisponde la relazione R di preordine, al sistema **S**5 una relazione R di equivalenza (ossia riflessiva simmetrica e transitiva). Leggermente più complicata risulta la situazione per i sistemi **S**1-**S**3; in questi casi infatti occorre introdurre un'ulteriore distinzione tra mondi *normali*, tali cioè che posseggano almeno un mondo a loro accessibile (in particolare se stessi), e mondi *non normali* per cui questa evenienza non si verifica.

L'analisi semantica che Kripke offriva per le logiche modali si inseriva in un generale ripensamento del significato di questi sistemi che aveva preso l'avvio fin dagli anni trenta con quella nota di Gödel di cui si è già parlato, in cui il sistema **S**4 di Lewis veniva assiomatizzato non ponendo più come centrale la nozione di implicazione stretta, ma considerandolo come un'estensione della logica proposizionale classica **CPC** in un linguaggio in cui figurava, oltre ai connettivi classici, l'operatore di necessità $\Box$. L'implicazione stretta veniva definita ponendo

$$\mathscr{A} \prec \mathscr{B} =_{\mathrm{Df}} \Box(\mathscr{A} \rightarrow \mathscr{B})$$

mentre per definizione $\Diamond \mathscr{A}$ era $\neg \Box \neg \mathscr{A}$.

In questa prospettiva, i sistemi di Lewis da calcoli dell'implicazione stretta si trasformavano in calcoli modali e in primo piano veniva il problema di caratterizzare le possibili nozioni di necessità. Seguendo l'impostazione di Gödel, nel 1957 Edward J. Lemmon forniva nuove assiomatizzazioni dei sistemi di Lewis, e di altri posteriori, come estensioni di **CPC** ottenute mediante l'aggiun-

zione di regole e assiomi specificamente modali. La distinzione di fondo era tra sistemi normali e non. I primi sono i sistemi chiusi rispetto alla regola di necessitazione

$$\text{RN} \qquad \frac{\mathscr{A}}{\Box\,\mathscr{A}},$$

la regola del modus ponens e che hanno tra i loro teoremi tutte le formule del tipo

$$(1) \qquad \Box(\mathscr{A} \rightarrow \mathscr{B}) \rightarrow (\Box\,\mathscr{A} \rightarrow \Box\,\mathscr{B}).$$

Il più debole dei sistemi normali già noti all'epoca della sistemazione di Lemmon era il sistema **T** (oppure **M**) di Feys e von Wright i cui assiomi modali sono tutte le formule di tipo (1) e quelle del tipo

$$\Box\,\mathscr{A} \rightarrow \mathscr{A}.$$

Da **T** si ottiene il sistema **S**4 di Lewis aggiungendo come assiomi le formule

$$\Box\,\mathscr{A} \rightarrow \Box\Box\,\mathscr{A}.$$

S5, a sua volta, si ottiene da **S**4 con l'aggiunta degli assiomi del tipo

$$\Diamond\,\mathscr{A} \rightarrow \Box\Diamond\,\mathscr{A}.$$

Lemmon indagava anche le possibili assiomatizzazioni dei sistemi non normali (tra i quali si trovano ad esempio **S**1, **S**2, **S**3, di Lewis) e introduceva nuovi sistemi sia normali che non normali tentando di fornire un'analisi dello spettro dei possibili sistemi modali sul piano sintattico. Molte altre ricerche furono condotte in quest'ottica, ad esempio indagando sistematicamente le riducibilità tra modalità presenti nei diversi sistemi, come aveva iniziato a fare William T. Parry già nel 1939. A queste analisi si affiancavano studi di carattere semantico condotti in termini di particolari arricchimenti delle algebre di Boole come già Tarski e McKinsey avevano fatto per **S**4 nel 1948. Fu ancora Lemmon che nel 1966 gettò le basi

per una indagine sistematica, dal punto di vista algebrico, dei vari calcoli modali, studiandone in quest'ottica, ad esempio, la decidibilità e i collegamenti con la semantica di Kripke.

Malgrado l'interesse di tutti questi metodi, è innegabile che il vero salto di qualità avvenne con l'impiego sistematico delle strutture di Kripke. Così ad esempio, come detto sopra, era possibile stabilire una vera e propria corrispondenza fra assiomi modali e proprietà della relazione di accessibilità, nel senso che sistemi con dati assiomi coincidevano con i calcoli caratterizzabili sul piano semantico in termini di strutture di Kripke con relazioni di accessibilità di un dato tipo: la transitività di R corrispondeva alla validità degli schemi di assiomi tipici di **S**4, la proprietà di essere un'equivalenza a quella degli assiomi **S**5, e così via.

Vedremo più avanti come negli anni sessanta siano emersi limiti di questa corrispondenza e in generale della semantica di Kripke nella classificazione dei sistemi modali; ciò che non toglie tuttavia che fu attraverso le strutture di Kripke che ricercatori come Krister Segerberg, Dov Gabbay, David Makinson e molti altri gettarono le basi di una vera e propria teoria dei modelli per la logica proposizionale modale studiando decidibilità, proprietà di interpolazione, classificando le estensioni e le riduzioni fra modalità, ecc.

L'analisi di Kripke portava d'altra parte ad una svolta decisiva nello sviluppo delle logiche intensionali nel loro complesso e forniva uno strumento unificante in grado di mettere in luce relazioni per l'innanzi insospettate, fra tutti i diversi sistemi noti portando alla creazione di nuovi. È difficile sopravvalutare il significato di questo risultato se si pensa che uno degli ostacoli centrali alla pacifica accettazione delle logiche modali (e in generale intensionali) è sempre stato la mancanza di una semantica chiara e sufficientemente articolata da permettere di ottenere risultati di completezza che il vecchio metodo delle matrici non era in grado di offrire. A partire dagli anni sessanta si è assistito così ad un poderoso sviluppo delle ricerche in questo campo; alla individuazione di strutture modello (le coppie $< I, R >$ sopra introdotte) in grado di fornire interpretazioni rispetto alle quali i vari sistemi conosciuti fossero completi, ha fatto seguito l'applicazione del concetto di sistemi di indici con relazione di accessibilità all'analisi di altre nozioni di carattere intensionale le cui connessioni con la nozione di modalità logica erano sino ad allora rimaste nel vago. Alludiamo

ai vari sistemi di *logica deontica* (nei quali compaiono operatori modali del tipo «è permesso», «è obbligatorio», ecc.), *di logica epistemica* (operatori «è noto che», «si crede che», ecc.) di *logica temporale* (nella quale gli operatori «sarà in futuro», «fu in passato» vengono trattati come «modalità») la cui connessione, presentata in una prospettiva unitaria da Georg von Wright già nel 1951, veniva così giustificata a posteriori.

Dal punto di vista filosofico il contributo più notevole dell'analisi di Kripke stava forse nel consentire di «riscattare» la logica modale dalle accuse di «essenzialismo» aristotelico che già dalla fine degli anni quaranta Quine le aveva ripetutamente lanciato. Le critiche di Quine colpivano quei tentativi che, a partire dai lavori di Ruth Barcan Marcus del 1946, erano stati condotti per estendere la logica modale enunciativa ad una predicativa del primo ordine. In tali sistemi infatti comparivano tesi del tipo di $\Diamond \exists xP \rightarrow \exists x \Diamond Px$, la cosiddetta *formula della Barcan*, il cui significato essenzialista è difficile negare, in quanto essa afferma come lecito il passaggio dalla possibile esistenza di un ente all'esistenza di un ente possibile. Analoghe difficoltà ancora più insidiose comparivano una volta che si volesse introdurre il concetto di identità e sviluppare su basi modali una logica del primo ordine con identità.[6] Risultava dubbia l'estensione della legge dell'indiscernibilità degli identici che sta alla base della stessa logica dell'identità. Orbene, Kripke riusciva nel 1963 a estendere la sua semantica al caso predicativo in modo tale da rendere impossibile la derivazione della formula della Barcan; in questo modo si apriva la strada alla costruzione di estensioni predicative dei vari sistemi modali, semanticamente complete rispetto alla interpretazione di Kripke e nello stesso tempo immuni dalle mende lamentate da Quine.

[6] È sufficiente per rendersi conto di questo che il lettore consideri l'esempio seguente: interpretiamo il simbolo «□» come «è logicamente necessario». Allora vale chiaramente

1) $\Box\ (12 > 7)$

d'altra parte è anche vero

2) il numero degli apostoli = 12.

Ricorrendo alle ordinarie proprietà delle identità otterremmo allora:

3) □ (il numero degli apostoli > 7).

La 3) è chiaramente assurda. Ne scende quindi che l'interpretazione intuitiva di «□» e le leggi dell'identità non sono compatibili.

L'idea di fondo è semplicemente quella di assegnare ad ogni mondo possibile un insieme di oggetti esistenti in esso; di operare quindi la riunione di tutti questi insiemi e di interpretare su di essa ogni costante predicativa del linguaggio; l'insieme U che si ottiene in questo modo rappresenta intuitivamente l'insieme di oggetti «possibili», ossia che esistono in qualche mondo. A questo punto l'interpretazione dei quantificatori non viene estesa a tutto U ma la verità in un dato mondo i di un enunciato contenente in quantificatore verrà valutata restringendo l'interpretazione del quantificatore stesso all'insieme degli oggetti di U che si trovano nel dato mondo i. È facile verificare che in questo modo la formula della Barcan non risulta vera in tutte le interpretazioni e si possono esibire controesempi per essa. Questo impone a livello sintattico che l'estensione predicativa del sistema modale di base non venga ottenuta trasportando direttamente gli assiomi e le regole classici, ma facendo opportune modifiche di cui Kripke forniva un primo esempio.

La semantica predicativa di Kripke non è l'unica possibile: altre ne furono sviluppate da Hintikka, R. Thomason e altri su linee analoghe e studi sistematici sulle varie alternative sono stati condotti da Kit Fine, Kenneth Bowen e altri negli anni settanta. Le variazioni riguardano sostanzialmente l'interpretazione delle costanti predicative e quella dei quantificatori. Discorso analogo si può fare per l'identità e lavorando opportunamente a livello semantico si è giunti a diverse formulazioni della logica modale con identità che, pur discostandosi dall'interpretazione classica, risultano facilmente assiomatizzabili. Va detto comunque che in questo campo molte questioni rimangono aperte e molte alternative devono essere ancora indagate; la cosa non sorprende se si pensa che questi problemi sono strettamente connessi ad antichi interrogativi filosofici sulla natura dell'identità, della necessità e del concetto di individuo.

Da quanto abbiamo detto, e dalla sopra ricordata traduzione di Gödel del 1933, risulta chiaro come sia possibile trovare una semantica adeguata in questo contesto per la logica intuizionista utilizzando strutture di Kripke definite su insiemi preordinati $<I,R>$ (dove cioè R è riflessiva e transitiva). Indicheremo con

$$\mathfrak{M} = <<I,R>,\mathrm{M}>$$

ogni struttura per cui, se iRj e $M_i \models \mathscr{A}$ allora $M_j \models \mathscr{A}$ per ogni atomica $\mathscr{A}$. Saranno queste le strutture di cui ci serviremo. La verità di una formula $\mathscr{A}$ a livello i in $\mathfrak{M}$ ($\mathfrak{M} \models_i \mathscr{A}$) viene definita per induzione sulla complessità di $\mathscr{A}$. Ci limiteremo a funzioni di interpretazioni M definite su insiemi preordinati $< I, R >$ per i quali valga che: per ogni $i, j \in I$, e per ogni *atomica* $\mathscr{A}$, se $\mathscr{A}$ è classicamente vera in i ($M_i \models \mathscr{A}$) e iRj allora $\mathscr{A}$ è vera classicamente in j ($M_j \models \mathscr{A}$).

In altre parole, una formula atomica, se vera, lo è necessariamente. A questo punto definiamo la nozione di verità in un livello i come segue:

1) per $\mathscr{A}$ atomica: $\mathfrak{M} \models_i \mathscr{A}$ se e solo se $M_i \models \mathscr{A}$;

2) $\mathfrak{M} \models_i \mathscr{A} \wedge \mathscr{B}$ se e solo se $\mathfrak{M} \models_i \mathscr{A}$ e $\mathfrak{M} \models_i \mathscr{B}$;

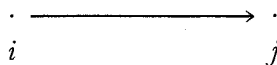
3) $\mathfrak{M} \models_i \mathscr{A} \vee \mathscr{B}$ se e solo se $\mathfrak{M} \models_i \mathscr{A}$ o $\mathfrak{M} \models_i \mathscr{B}$;

4) $\mathfrak{M} \models_i \mathscr{A} \rightarrow \mathscr{B}$ se e solo se per ogni j tale che iRj non $\mathfrak{M} \models_j \mathscr{A}$ o $\mathfrak{M} \models_j \mathscr{B}$;

5) $\mathfrak{M} \models_i \neg \mathscr{A}$ se e solo se per ogni j tale che iRj non $\mathfrak{M} \models_j \mathscr{A}$.

Come si nota le clausole di verità colgono appieno il senso dell'interpretazione intuizionista; se interpretiamo infatti i mondi possibili come diversi stati di evidenza (diversi momenti nella costruzione di una «dimostrazione») e R come la relazione di passaggio da uno di questi stati all'altro, risulta ad esempio – limitandoci per comodità al caso cruciale della negazione – che $\neg \mathscr{A}$ è vera se e solo se non potrà mai raggiungersi uno stato di evidenza per $\mathscr{A}$ (in altri termini, se $\mathscr{A}$ è assurda).

Si vede così come è possibile refutare il terzo escluso: basterà considerare una struttura a due livelli

$$\underset{i}{\cdot} \longrightarrow \underset{j}{\cdot}$$

dove a livello i non è vera $\mathscr{A}$ che invece è vera a livello j. È chiaro che non potrà essere $\mathfrak{M} \models_i \mathscr{A} \vee \neg \mathscr{A}$ in quanto per ipotesi non $\mathfrak{M} \models_i \mathscr{A}$ e non può essere $\mathfrak{M} \models_i \neg \mathscr{A}$ in quanto iRj e $\mathfrak{M} \models_j \mathscr{A}$.

Per quanto riguarda i quantificatori, assegneremo ad ogni mondo i un insieme di oggetti $\varphi(i)$ con la clausola che, se iRj, $\varphi(i) \subseteq$

$\varphi(j)$, e interpreteremo ogni lettera predicativa n-adica, a seconda dello stato di evidenza in cui ci troviamo, come un sottoinsieme di $\varphi(i)^n$. Porremo allora che $\mathfrak{M} \models_i \exists x \mathscr{A}$ se esiste un individuo $a \in \varphi(i)$ tale che $\mathfrak{M} \models_i \mathscr{A}(a)$; per converso, nel caso di una formula universale, $\forall x \mathscr{A}$, porremo che $\mathfrak{M} \models_i \forall x \mathscr{A}$ se e solo se per ogni $j \in I$ tale che iRj , si ha che $\mathfrak{M} \models_j \mathscr{A}(a)$ per ogni $a \in \varphi(j)$.

Come ricordato all'inizio, e come è facile vedere, l'interpretazione di Kripke non è costruttiva come pure non costruttiva è la dimostrazione di completezza che egli dava rispetto a questa semantica, utilizzando metodi introdotti da Beth. Allo stesso Beth si deve d'altra parte il primo tentativo di semantica non algebrica per la logica intuizionista che risale al 1956 e che tuttoggi ha grande interesse proprio in quanto più dominabile dal punto di vista costruttivo.

Beth partiva dai concetti di albero e di sbarramento e considera strutture $\mathfrak{B}$ analoghe a quelle di Kripke, con l'unica vincolante appunto che $< I, R >$ è un albero con origine α e non un preordine qualunque. Ogni formula viene verificata livello per livello e le clàusole che Beth dava differiscono da quelle di Kripke da due punti di vista: in primo luogo, diversamente da Kripke, Beth assume un *unico* dominio prefissato D per ogni livello; in secondo luogo utilizza il concetto di sbarramento per definire la verità di una formula $\mathscr{A}$ a livello i nel caso delle formule atomiche, delle disgiunzioni e dell'esistenziale. Per il resto la situazione è identica a quella considerata da Kripke. Così avremo che se $\mathscr{A}$ è atomica $\mathfrak{B} \models_i \mathscr{A}$ se esiste un insieme X che sbarra i, tale che per ogni $j \in X$, $\mathfrak{B} \models_j \mathscr{A}$ dove dire che X sbarra i significa, come si ricorderà, che ogni ramo che passa per i interseca X.

Analogamente per $\vee$ e $\exists$:

$\mathfrak{B} \models_i \mathscr{A} \vee \mathscr{B}$ sse esiste un X che sbarra i tale che per ogni $j \in X$, $\mathfrak{B} \models_j \mathscr{A}$ oppure $\mathfrak{B} \models_j \mathscr{B}$; avremo invece $\mathfrak{B} \models_i \exists x \mathscr{A}$ se e solo se esiste un X che sbarra i tale che per ogni $j \in X$ esiste un a_j per cui $\mathfrak{B} \models_j \mathscr{A}(a_j)$.

Diciamo infine che una formula $\mathscr{A}$ è vera in $\mathfrak{B}$ ($\mathfrak{B} \models \mathscr{A}$) quando è vera, nel senso precedente, nell'origine α. È facile verificare che la logica intuizionista è valida rispetto all'interpretazione di Beth, mentre la completezza richiede ancora una volta un ricorso al lemma di König e, quindi, al teorema di sbarramento di Brouwer. Questo ci mostra l'origine dell'idea intuitiva di Beth: ogni interpretazione rappresenta un possibile sviluppo della conoscenza

di un soggetto e dire che un livello i è sbarrato da un insieme X significa che in qualunque modo si proceda da i si dovrà comunque passare per X.

Il significato dei lavori di Kripke e di Beth non si ferma però qui e nel 1970 K. Segerberg è riuscito a estendere la semantica in questione a sistemi ancora più deboli di quello intuizionista, quale ad esempio la logica minimale, e ad analizzare tutto il segmento di sistemi intermedi che esistono fra questa e quella classica; da allora a oggi le ricerche sulle logiche intermedie che utilizzano i metodi di Kripke, o tecniche connesse, si sono moltiplicate.

Fra i vari sistemi non classici legati strettamente alle logiche modali e che si sono giovati della sistemazione della semantica basata sull'introduzione dei «mondi possibili» (logiche deontiche, logiche epistemiche) vogliamo qui soffermarci sulla cosiddetta *logica del tempo* (detta anche *logica cronologica*) la cui nascita – ma è più esatto dire la rinascita[7] – ha coinciso – ed è particolarmente interessante sottolinearlo in questa sede – con l'approfondimento, intrapreso soprattutto da Arthur Norman Prior (1917-69) – di

[7] È noto che già Aristotele nel *De interpretatione* escludeva che il principio del terzo escluso valesse per gli enunciati contingenti (non necessari) al futuro, ponendo così le basi della logica trivalente formalizzata da Łukasiewicz. Nella logica medievale, araba ed europea, non è mai venuta meno la convinzione che le proposizioni potessero cambiare valore di verità con il tempo e che il tempo fosse un oggetto pertinente alla speculazione logica. Con la nascita della logica contemporanea, invece, nonostante le osservazioni in contrario di filosofi come MacTaggart e Peirce, e di logici come J.N. Findlay (autore, nel 1941, di un articolo in cui si proponeva l'introduzione di operatori temporali) Reichenbach e soprattutto Łos, vero fondatore a tutti gli effetti della logica cronologica (1949), hanno prevalso le tesi a favore della riducibilità degli enunciati contenenti forme verbali temporali ad enunciati contenenti una copula atemporale. Su questo atteggiamento ha pesato indubbiamente la considerazione privilegiata degli enunciati della matematica, chiaramente atemporali. Conviene qui accennare che la tesi più drasticamente polemica nei confronti della opportunità di una logica del tempo è quella sostenuta da Quine e da J.C. Smart, per i quali ogni enunciato è riducibile ad un enunciato atemporale in cui la proposizione temporale è convertibile in un predicato relativo a date (p. es. «Cesare morì nel 44 a.C.» è traducibile in «è atemporalmente vero dell'anno 44 a.C. che Cesare morì in questa data» oppure «l'anno 44 a.C. è [atemporalmente] caratterizzato dalla morte di Cesare»). La risposta di Prior, così come di altri studiosi, a queste obiezioni è naturalmente che il tempo verbale è ineliminabile dagli enunciati: per fare un solo esempio, l'interpretazione Quine-Smart appiattisce per Prior la distinzione da lui ritenuta fondamentale tra «non era vero nel 44 a.C. che p» ed «era vero nel 44 a.C. che non p».

studi di storia della logica. Supponendo di dover interpretare infatti l'operatore modale di possibilità in termini temporali come «è o sarà vero che», ci si ritrova con la concezione del possibile professata da Diodoro Crono, un megarico del III secolo a.C. Gli studi su Diodoro suscitati da queste osservazioni storiche hanno mirato da un lato a ricostruire la forma esatta del cosiddetto «argomento vittorioso» di Diodoro (secondo il quale, partendo dalle premesse che il passato è necessariamente passato e che ciò che implica l'impossibile è esso stesso impossibile, Diodoro giungeva appunto a identificare il *possibile* con *ciò che è o sarà vero*); dall'altro hanno condotto a tentare di stabilire quale dei sistemi modali poteva costituire la corretta interpretazione della nozione diodorea di possibilità. L'individuazione del sistema diodoreo è stato il punto più difficile da determinare. La prima scoperta di Prior fu che il comportamento della modalità diodorea era rispecchiato da una matrice infinita che verificava tutte le tesi di **S**4 ma non tutte quelle di **S**5. Prior concludeva quindi un po' precipitosamente, in *Time and modality* (*Tempo e modalità*, 1957), che la matrice era caratteristica per **S**4 e che di conseguenza il sistema diodoreo corretto era **S**4. Non tardarono a farsi sentire le critiche di Lemmon, Hintikka e Kripke i quali indipendentemente attorno al 1957-1958 presentarono una formula $(L(Lp \rightarrow Lq) \vee L(Lq \rightarrow Lp)$ nella formulazione di Lemmon, ove L sta per $\square$) che pur essendo diodorea non era in **S**4: sembrava dunque che **S**4.3 (= **S**4 + la formula di Lemmon) dovesse essere il sistema cercato. La ricerca si è conclusa invece soltanto aggiungendo ad **S**4.3 la formula (dove M sta per $\diamond$) $L(L(p \rightarrow Lp) \rightarrow p) \rightarrow (MLp \rightarrow p)$, che M. Dummett nel 1958 dimostrava essere soddisfatta dalla matrice di Prior e purtuttavia non derivabile in **S**4.3.

Ciò che Prior intendeva precisamente come logica del tempo era chiarito in questi termini da Prior stesso: logica del tempo è una logica che *a*) contiene variabili enunciative che variano su enunciati che siano veri in taluni momenti e falsi in altri; *b*) contiene le normali funzioni di verità definite con opportune modifiche, e cioè in modo che p sia vera *quando* e solo quando sia falsa p, ecc.; *c*) contiene funzioni addizionali come Fp e Pp interpretabili «sarà vero che p» ed «è stato vero che p», nonché funzioni interdefinibili con queste come $Gp = \neg F \neg p$ (sarà sempre vero che p) e $Hp = \neg P \neg p$ (è stato sempre vero che p).

Un modo diverso di fondare una logica temporale è quello (preferito da logici come G.H. von Wright, G.E.M. Anscombe, D.R. Luce) di introdurre un connettivo primitivo, diversamente caratterizzato, esprimente la relazione «prima di». Per illustrare la differenza, diamo un esempio di base assiomatica per i due tipi di logica modale: il sistema minimale $\mathbf{K}_t$ di Lemmon (1959), e il sistema «And then» di von Wright (1966), che è stato appropriatamente chiamato «logica del cambiamento» ed è contraddistinto dal connettivo primitivo T.

$\mathbf{K}_t$: (con F e P indefiniti). Definizione di G: G = $\neg$ F $\neg$; Definizione di H: H = $\neg$ P $\neg$;

Regole: RG: $\mathscr{A} / \neg \mathrm{F} \neg \mathscr{A}$; RH: $\mathscr{A} / \neg \mathrm{P} \neg \mathscr{A}$;

Assiomi:
1.1 $\neg \mathrm{F} \neg (p \rightarrow q) \rightarrow (\mathrm{F}p \rightarrow \mathrm{F}q)$;
1.2 $\neg \mathrm{P} \neg (p \rightarrow q) \rightarrow (\mathrm{P}p \rightarrow \mathrm{P}q)$;
2.1 $\mathrm{P} \neg \mathrm{F} \neg p \rightarrow p$;
2.2 $\mathrm{F} \neg \mathrm{P} \neg p \rightarrow p$.

«And then». Regole: sostituzione, *modus ponens*, estensionalità.

Assiomi:
A_t0. gli assiomi del calcolo classico delle proposizioni;
A_t1. $((p \vee q) \mathrm{T} (r \vee s)) \longleftrightarrow ((p \mathrm{T} r) \vee (p \mathrm{T} s) \vee (q \mathrm{T} r) \vee (q \mathrm{T} s))$;
A_t2. $((p \mathrm{T} q) \wedge (p \mathrm{T} r)) \longleftrightarrow (p \mathrm{T}(q \wedge r) \vee (q \mathrm{T} r) \vee (r \mathrm{T} q))$;
A_t3. $p \longleftrightarrow (p \mathrm{T} (q \vee \neg q))$;
A_t4. $\neg (p \mathrm{T} (q \wedge \neg q))$.

Può comunque essere fuorviante insistere su questa distinzione per avere un criterio di orientamento nel problema della logica del tempo: fuorviante nella misura in cui i due tipi di sistema risultano in linea di principio intertraducibili e nella misura in cui i sistemi di logica del tempo sono classificabili con criteri anche più profondi. Alludiamo alla distinzione (da Lemmon detta «stratificazione») dei postulati temporali in postulati esprimenti le condizioni di verità di enunciati temporalizzati, quali sono quelli inclusi nel sistema minimale $\mathbf{K}_t$, e postulati, addizionabili a $\mathbf{K}_t$, che riflettono proprietà della relazione prima-dopo. Ci sono così assiomi

che presuppongono la densità e la discretezza dell'ordine temporale, altri che esprimono la circolarità, o l'infinità, o la finitezza della serie temporale, altri ancora che rispecchiano la linearità, o la ramificazione del futuro ecc. Inutile sottolineare che aggiungendo coerentemente assiomi di questo secondo «strato» al sistema minimale si generano sistemi differenziati a livello semantico (la relazione di accessibilità tra diversi stati del mondo è nella logica cronologica una relazione temporale).

Un diverso ordine di differenziazioni si apriva passando dalla logica temporale enunciativa a quella predicativa, in particolare per quanto riguarda i delicati rapporti tra esistenza e quantificazione, che fanno entrare in gioco considerazioni ontologiche altrettanto importanti di quelle sulla struttura del tempo. Se per esempio si interpreta, come faceva abitualmente Prior, il quantificatore esistenziale come tale da variare sugli enti *attualmente esistenti*, la formula della Barcan risulta chiaramente falsa in quanto esprime un passaggio non garantito dalla possibilità all'esistenza. Lo stesso ragionamento vale sotto l'interpretazione temporale dell'operatore di possibilità abitualmente data da Prior. Il problema ontologico dunque si ripresentava con il consueto interrogativo: quali sono gli enti di cui si intende ammettere l'esistenza? Le risposte a questo problema sono state svariate. La formula della Barcan risulta valida per esempio nell'interpretazione temporale del tipo Quine-Smart, dove «Mp» significa «$\exists tpt$» (dove p è la proposizione predicato) e anche in alcuni sistemi studiati da Prior in *Tempo e modalità* ($\Sigma\mathbf{T}_2$ e $\Sigma\mathbf{T}_3$). Alla restrittività delle assunzioni ontologiche di Prior faceva contrasto la liberalità con cui Nino B. Cocchiarella (autore, nel 1966, di una fondamentale tesi di dottorato sulla logica del tempo) accoglie nel suo universo tanto gli enti attuali che gli enti possibili, introducendo due tipi di quantificatori esistenziali, ammettendo così entrambe le interpretazioni.

Questo sviluppo impetuoso delle logiche intensionali a partire dalla fine degli anni cinquanta contrasta in qualche modo con la lentezza che ha segnato lo svolgersi delle indagini sulle logiche polivalenti, che al loro apparire erano sembrate a molti il prototipo delle logiche non classiche.

Come si ricorderà, nell'articolo di Łukasiewicz e Tarski del 1930 *Ricerche sul calcolo proposizionale* venivano presentati una serie di calcoli L_n (dove n è un qualunque numero naturale o $n = \aleph_0$,

che indica appunto il numero di «valori di verità» assunti nel calcolo stesso). In particolare tali valori di verità venivano «stardardizzati» assumendoli uguali a $k/(n-1)$ (con $0 \leqslant k \leqslant n-1$) nel caso di $1 < n < \aleph_0$; e uguali a k/l (con $0 \leqslant k < l$) nel caso di $n = \aleph_0$. Le «tavole di verità» (matrici) per i connettivi di implicazione e negazione venivano definite ponendo (se x e y sono due valori individuati come sopra) $p \rightarrow q = 1$ se x (valore di verità di p) è minore o uguale a y (valore di verità di q); $p \rightarrow q = 1 - x + y$, se $x > y$; $\neg p = 1 - x$.[8] Si è già ricordato che in questo articolo venivano dimostrati tutta una serie di risultati *metalogici* su questi calcoli, ad esempio che $L_m \subseteq L_n$ se e solo se $n-1$ divide $m-1$ (Lindenbaum). Dopo che Wajsberg aveva assiomatizzato il calcolo trivalente sulla base del sistema di assiomi

1) $p \rightarrow (q \rightarrow p)$
2) $(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow (p \rightarrow r))$
3) $(\neg p \rightarrow \neg q) \rightarrow (q \rightarrow p)$
4) $((p \rightarrow \neg p) \rightarrow p) \rightarrow p$

con le regole del *modus ponens* e di sostituzione, la cosa venne generalizzata per ogni n strettamente *minore* di $\aleph_0$ da Lindenbaum che estese il risultato di Wajsberg dimostrando che ogni sistema con un numero finito di valori di verità è assiomatizzabile. Nello stesso articolo Łukasiewicz congetturava che anche il sistema $L_{\aleph_0}$ fosse assiomatizza-

[8] Supponiamo ad esempio di considerare L_5; allora i valori di verità saranno 0,1/4,2/4 (ossia 1/2), 3/4, 4/4 (ossia 1). La tavola di verità per l'implicazione sarà

→	0	1/4	1/2	3/4	1
0	1	1	1	1	1
1/4	3/4	1	1	1	1
1/2	1/2	3/4	1	1	1
3/4	1/4	1/2	3/4	1	1
1	0	1/4	1/2	3/4	1

e quella per la negazione

p	$\neg p$
0	1
1/4	3/4
1/2	1/2
3/4	1/4
1	0

si noti che nel caso di logiche trivalenti si usa indicare i tre valori con 0,1/2,1, oppure con 0,1,2.

bile sulla base degli assiomi 1), 2), 3) come sopra, con l'aggiunta di

4) $((p \rightarrow q) \rightarrow q) \rightarrow ((q \rightarrow p) \rightarrow p)$
5) $((p \rightarrow q) \rightarrow (q \rightarrow p)) \rightarrow (q \rightarrow p)$

con le solite regole di sostituzione e del *modus ponens*.[9] Nel 1936 Jerzy Słupecki in *Der volle dreiwertige Aussagenkalkül* (*Il calcolo proposizionale trivalente pieno*) dimostrava che il calcolo trivalente di Łukasiewicz come assiomatizzato da Wajsberg non era *funzionalmente completo*, cioè non era possibile costruire, per mezzo dei connettivi primitivi in esso assunti, *tutte* le funzioni proposizionali (con un numero finito di argomenti) e ovviava alla cosa con l'introduzione di un nuovo operatore *T* (il cosiddetto «tertium» di Słupecki) e di due nuovi assiomi che ne regolano il comportamento. Su un altro versante, nel 1939 D.A. Bochvar in *Sul calcolo logico polivalente e la sua applicazione all'analisi delle contraddizioni*, presentava una variante del calcolo trivalente di Łukasiewicz con l'obiettivo di affrontare le antinomie della teoria degli insiemi ricorrendo appunto ad una logica che ammettesse tre valori di verità. Allo stesso anno risale il primo tentativo, di John Barkley Rosser (1907-89), di considerare un calcolo *predicativo* polivalente, ma solo nel 1958 se ne dà una soddisfacente presentazione assiomatica nel volume dello stesso Rosser e di Atwell R. Turquette, *Many-valued Logics* (*Logiche polivalenti*); cui seguirono altri interessanti contributi di Rosser (1960), di Mostowski (1961) e Chang (1958, 1964).

Fu Chang che nel 1958 compì il passo decisivo analizzando la struttura delle algebre che sottostanno alle logiche polivalenti come quelle di Boole alla logica classica o quelle di Heyting all'intuizionista. Queste algebre – note appunto come *MV-algebre* – hanno una struttura complessa e studiando la generalizzazione del concetto di filtro per esse, Chang riuscì a ottenere risultati di completezza per i calcoli presentati, risultati che saranno riformulati successivamente, nel 1974, da Dana Scott, ponendone in luce contatti con gruppi abeliani e moduli. Ma la difficoltà centrale ri-

[9] L'assioma 4) è stato dimostrato essere superfluo (ossia dipendere dai rimanenti) indipendentemente da Chen Chung Chang e da C.A. Meredith nel 1958. La congettura di Łukasiewicz è stata dimostrata corretta nel 1958 da Alan Rose e Rosser nell'articolo *Fragments of many-valued statement calculi* (*Frammenti di calcoli proposizionali polivalenti*).

maneva quella di trovare una giustificazione *filosofica* o *matematica* plausibile per queste logiche, al di là della possibilità di condurne un'analisi algebrica. Né a suo tempo Łukasiewicz, né più tardi altri che ci provarono, riuscirono a dare un senso alle definizioni di Łukasiewicz che non ricorra a semplici analogie col caso classico o a dubbie interpretazioni di tipo probabilistico. I tentativi di Scott hanno mostrato come collegare logiche polivalenti a strutture tipo Kripke, utilizzando una interpretazione delle formule in termini di «approssimazione» (come ad esempio – e qui è immediato il legame con l'analisi della continuità di Poincaré – quando si valuta l'identità di punti su una retta) mentre i risultati di Chang del 1965 che riprendevano l'idea di Bochvar di analizzare i paradossi utilizzando più valori di verità, sono rimasti un vicolo cieco, come del resto precedenti tentativi al riguardo di Skolem del 1957.

Rimaneva l'aspetto astratto e la possibilità – realizzata da Chang in collaborazione con Keisler in un libro del 1966 dal titolo *Continous Model Theory* (*Teoria continua dei modelli*) – di estendere molti dei risultati classici di teoria dei modelli a logiche polivalenti, comprese appunto quelle continue, ma nel complesso non si può dire che la ricerca degli anni sessanta abbia portato a risultati paragonabili a quelli ottenuti per le logiche modale e intuizionista.

L'aggancio più profondo tra polivalenza e concetto di insieme sarebbe venuto però dalla stessa teoria degli insiemi e non da speculazioni generiche e pesantemente connotate ideologicamente, come quelle di Łukasiewicz. Qui non è tanto l'idea di polivalenza come indeterminazione che contò (Łukasiewicz vedeva il rigetto della bivalenza come rifiuto del determinismo) quanto l'idea di variazione ad essere il punto centrale. Dopo che nel 1963 J. Paul Cohen aveva dimostrato col metodo di *forcing* (di cui parleremo nel prossimo paragrafo) l'indipendenza dell'assioma di scelta e dell'ipotesi generalizzata del continuo dagli altri assiomi della teoria degli insiemi, Dana Scott e Robert Solovay nel 1967 offrirono quella che a nostro parere rimane la più significativa applicazione metamatematica dell'idea di polivalenza riproponendo la dimostrazione del risultato di Cohen sulla base di un'interpretazione polivalente della teoria degli insiemi, dove i «valori di verità» sono elementi di un'opportuna algebra di Boole.

Ma di questo ci occuperemo più avanti parlando della teoria degli insiemi. Qui vogliamo tornare sul tema delle logiche non classi-

che e ai loro sviluppi dagli anni sessanta, in particolare sulle logiche intensionali che ebbero in questo contesto un ruolo di guida.

Negli anni sessanta e nei primi anni settanta, le ricerche nell'ambito della logica dell'intensione mantennero infatti lo slancio inizialmente ricevuto dalle scoperte di Saul Kripke e di quanti avevano contribuito, intorno al 1960, a elaborare la semantica modale. Prova ne sia la pubblicazione di manuali dedicati alla logica modale quali quello di Hughes e Cresswell del 1968 e di J. Zeeman del 1973. Si moltiplicavano intanto le applicazioni nuove delle tecniche di Kripke a settori della logica fino ad allora inesplorati, come pure quelle relative ad altre discipline (linguistica, informatica, la stessa matematica).

Questa tendenza era accompagnata all'altra, altrettanto viva, a non disperdersi nella ricerca di risultati parziali relativi a singoli sistemi, per tentare uno studio organico e generale della pluralità (sempre più difficile da dominare) dei possibili sistemi modali. Questa esigenza compare non solo in ricerche di ampio respiro come i tre volumi di Krister Segerberg *An Essay in classical modal logic* (*Un saggio di logica modale classica*, 1971) dedicati ad una esplorazione sistematica della semantica della logica modale classica, lungo la linea iniziata con i lavori di Lemmon ma anche in contributi più limitati come quelli di Bengt Hansson e Peter Gärdenfors del 1973 in cui gli autori cercavano di mettere ordine nello spettro dei sistemi modali imparentati ai sistemi di Lewis analizzando la potenza della semantica di Kripke in relazione alla cosiddetta «semantica di intorni» (*neighbourhood semantics*) elaborata da Scott e Richard Merritt Montague (1931-71) e alla semantica booleana. I due autori tentano un'analisi dei sistemi modali utilizzando le nozioni di *ampiezza* e di *profondità* di una struttura modale: *ampiezza* viene detta la misura dell'intervallo tra la logica più debole e la logica più forte tra quelle che possono essere determinate[10] da almeno una struttura del tipo in questione; *profondità* è detta invece la misura del numero delle logiche tra i due estremi che possono essere determinate da tali strutture. La nozione di misura che qui interviene non va tuttavia intesa in senso metrico e quantitativo; si tratta per-

[10] Si dice che una logica è *determinata* da un tipo di strutture quando i suoi teoremi coincidono con l'insieme delle formule valide rispetto alle interpretazioni che hanno come supporto strutture del tipo dato. Ad esempio il sistema **S4** di Lewis è determinato dalle strutture di Kripke $< X, R >$ dove R è riflessiva e transitiva.

tanto, come gli stessi autori ammettono, più di un programma di lavoro che di una ricerca sistematica, anche se entro questi limiti venivano dimostrati risultati molto interessanti, per esempio che la formulazione abituale della semantica di intorni è equivalente a una trattazione estensionale degli operatori modali. La domanda che diveniva sempre più pressante, in vista del successo della semantica di Kripke nell'organizzare la logica modale, era quella di sapere entro che limiti in generale si possano riflettere le proprietà delle relazioni di accessibilità dell'interpretazione di Kripke in termini di leggi e regole modali. Quello che emergeva era un blocco di problemi connessi alla traducibilità delle logiche modali nella logica classica del secondo ordine e quindi alla caratterizzabilità di classi di strutture del tipo $< I, R >$ mediante formule di data forma.

Risultati di estremo interesse in questa direzione non tardarono a venire. I più importanti sono probabilmente i teoremi di incompletezza dimostrati indipendentemente da Kit Fine e Steven K. Thomason nel 1974. Si è così provato che esistono sistemi modali che sfuggono alla semantica a mondi possibili nel senso che non sono determinati da classi di strutture-modello e le cui tesi non sono caratterizzabili come tutte e sole le formule del linguaggio modale valide in ogni interpretazione di Kripke definita su struttura $< I, R >$ che soddisfano specifiche proprietà. Risultati del genere si estendevano anche – come dimostrato da Martin Gerson – alle semantiche d'intorni in cui ad ogni mondo $i \in I$ è assegnata una famiglia Φ_i di insiemi di mondi e $\mathfrak{M} \models_i \Box \mathscr{A}$ se e solo se esiste un $X \in \Phi_i$ tale che per tutti i $j \in X$, $\mathfrak{M} \models_j \mathscr{A}$.

Il problema di quale sia il modo più generale di fornire semantiche per le logiche modali rimane tutt'oggi aperto, anche se il complesso di risultati specifici è imponente. Di particolare interesse per quanto riguarda il problema generale sono le ricerche di Robert Goldblatt e S.K. Thomason sulle *semantiche al secondo ordine* in cui si assume che ogni modello sia corredato di una famiglia privilegiata di insiemi di mondi, ciascuno dei quali corrisponde intuitivamente a una proposizione. La validità viene definita solo in riferimento alle interpretazioni che assegnano come possibili valori alle formule atomiche proposizioni come quelle di sopra, proprio come nei modelli generali di Henkin si sostituisce la vincolante a tutte le parti dell'universo con quella più debole che riguarda una determinata sottofamiglia di queste. È così possibile stabilire si-

stematicamente una dualità tra le semantiche a mondi possibili (sia nella forma di Kripke, che in quella di Montague, che in quella del secondo ordine) e le semantiche algebriche in termini di algebre di Boole con operatori.

Questi risultati riguardano la logica proposizionale, ma i dibattiti filosofici più rilevanti sulla modalità hanno in generale coinvolto gli sviluppi al primo ordine e agli ordini superiori in cui $\Box$ interagisce con identità e quantificatori. Già abbiamo parlato del carattere essenzialista di tesi come quella della Barcan (per cui dalla possibilità di esistenza si inferisce l'esistenza di un possibile) ampiamente criticata negli anni quaranta e cinquanta da Quine, e abbiamo ricordato anche come in molti sistemi sembri inevitabile inferire l'identità necessaria dall'identità a meno di rinunciare alla legge di Leibniz. Gli anni sessanta hanno visto interessanti indagini sulle alternative che si aprono in questo campo, in particolare sulla possibilità di eliminare dai teoremi la legge di Barcan e il passaggio dall'identità all'identità necessaria.

Questo comportava un'analisi sistematica dei possibili modi di assegnare domini ai singoli mondi che possiamo far iniziare con l'articolo di Kripke del 1963 *Semantical considerations on modal logics* (*Considerazioni semantiche sulle logiche modali*). Come si ricorderà, Kripke mostrava come la refutazione della legge di Barcan fosse connessa alla possibilità che, passando da un mondo ad uno accessibile, l'universo degli individui si allarghi e come la necessità di ogni identità corrispondesse al fatto che passando da un mondo ad uno accessibile si assumesse la conservazione tanto dell'esistenza quanto della diversità degli individui. Egli mostrava come, anche nel caso di sistemi con relazione di accessibilità simmetrica, fosse possibile, modificando la nozione di validità, refutare la formula della Barcan e in anni successivi, in un lavoro rimasto famoso dal titolo *Naming and Necessity* (*Nome e necessità*), pubblicato in volume nel 1972, introduceva il concetto di *designazione rigida*, la designazione che si ha quando un termine denota in ogni mondo lo stesso individuo, rendendo così necessaria l'identità che eventualmente lo coinvolge. Kripke portava argomenti di carattere filosofico per la validità di questa assunzione che dettero origine a accesi dibattiti tutt'oggi vivaci e che in alcuni casi si concretarono anche in esplicite formulazioni di carattere tecnico per la costruzione di semantiche del primo ordine, soprattutto per opera di Fi-

ne e Thomason. Un passo ulteriore verso il potenziamento della logica modale e in generale di quella intensionale, venne compiuto nel 1974 da un allievo di Montague, D. Gallin, nel volume *Intensional and higher order modal logic* (*Logica intensionale e modale di ordine superiore*), in cui sviluppava una vera e propria teoria degli oggetti intensionali (proposizioni, intensioni, ecc.) sotto forma di una logica modale di ordine superiore, parallela a quella classica dei *Principia.*

Queste ricerche nascevano al punto di confluenza di due linee di sviluppo, particolarmente vive: le indagini sulla semantica denotazionale dei linguaggi naturali (che già Montague aveva iniziato negli anni sessanta) e le ricerche sulla possibilità di sviluppare una vera e propria teoria modale degli insiemi. Il fatto interessante da osservare è che questa seconda direzione avrebbe acquisito particolare rilievo una volta che, negli anni settanta, si vide come i modelli booleani di Scott e Solovay, il forcing di Cohen e, appunto, sviluppi come quelli di Gallin potevano trovare un contesto unitario nell'interpretazione della logica superiore in categorie di fasci.

Un altro aspetto della ricerca logica di carattere intensionale particolarmente attivo a partire dagli anni sessanta è quello che riguarda le diverse forme di implicazione forte. Verso la fine degli anni sessanta, l'intenso lavorio legato al concetto di implicazione rilevante culmina – soprattutto per opera di Alasdair Urquhart e Richard Routley – in una trattazione semantica adeguata della logica dell'implicitazione (*entailment*) mediante l'impiego di una relazione di accessibilità ternaria. Come è noto il problema dell'*entailment* è quello di fornire un'analisi semantica e sintattica della relazione di implicazione in uso nel linguaggio comune. Il problema non è linguistico, ma ha come obiettivo piuttosto l'individuazione del nucleo minimale di proprietà che ogni nozione di implicazione deve avere. L'idea di fondo è che un enunciato del tipo $\mathscr{A} \supset \mathscr{B}$ deve essere vero solo quando $\mathscr{A}$ è *rilevante* nella deduzione di $\mathscr{B}$, nel senso che $\mathscr{A}$ deve essere usato *effettivamente* in qualche deduzione di $\mathscr{B}$.

I vari sistemi di implicazione stretta presentati allo scopo da Lewis a partire dal 1918 – e basati come si è visto sul concetto di possibilità – partivano dall'idea che perché $\mathscr{A}$ implichi strettamente $\mathscr{B}$ deve esistere un legame concettuale fra i due enunciati. Come si sa però nessuno dei sistemi di Lewis è completamente adeguato allo scopo: si ripresentano in essi – questa volta ovviamente per l'implicazione stretta – paradossi analoghi a quelli del-

l'implicazione materiale dei calcoli classici. Discorso analogo vale per l'implicazione nei calcoli intuizionisti che è ancora troppo «debole» in questo senso. Fu questa situazione che spinse prima Church nel 1951 e poi Ackermann nel 1956 a formulare calcoli formali per una *implicazione forte.* Queste ricerche culminarono con i sistemi per una *implicazione rilevante* presentati da Alan Ross Anderson e Nuel D. Belnap jr. a partire dal 1958, e che si basavano su modificazioni del sistema di Ackermann. Queste logiche ponevano come centrale il collegamento fra dimostrabilità e implicazione sopra indicato, sicché per esempio lasciavano cadere lo schema del sillogismo disgiuntivo, ossia lo schema $(((\mathscr{A} \vee \mathscr{B}) \wedge \neg \mathscr{A}) \supset \mathscr{B})$.

La debolezza dei sistemi di logica rilevante stava nella difficoltà di fornire un'analisi semantica della nozione presentata cosicché la loro giustificazione risultava spesso *ad hoc.* Un primo tentativo per ovviare a questa difficoltà si ebbe nel 1970 ad opera di J. Michael Dunn che utilizzò come modelli particolari strutture algebriche; se questo però forniva un metodo algebrico d'analisi dei sistemi, non era d'altra parte sufficiente per giustificare sul piano semantico le assunzioni di base. I risultati di Urquhart e Routley di cui sopra si è detto ne forniscono invece un'analisi plausibile in termini di strutture di Kripke generalizzate. Dal punto di vista filosofico non ci sembra però di poter dire che questi risultati rappresentino una giustificazione del tutto soddisfacente in quanto non è agevole collegare la semantica con l'analisi intuitiva in termini di dimostrabilità. In questa seconda direzione si muovono piuttosto alcuni lavori di Scott che partono dai calcoli di deduzione naturale e da un'analisi generale del loro significato, ma che però – dopo aver fornito strumenti estremamente eleganti per una teoria generale delle relazioni di conseguenza – non sono stati poi ulteriormente sviluppati.

Questa ripresa degli studi semantici sull'implicitazione logica coincideva con un analogo intensificarsi di studi sull'implicitazione fisica (condizionali congiuntivi e controfattuali, condizionali indicativi forti) in cui i risultati più significativi sono stati ottenuti da Robert Stalnaker e Richmond Thomason nel 1970, da Lennart Åqvist nel 1971 e da David K. Lewis nel 1973. Il problema dei controfattuali dipende, come è noto, dalla loro ambiguità contestuale, per cui essi vengono asseriti sulla base di presupposizioni non esplicitate in forza delle quali talora una stessa ipotesi rende plausibili con-

seguenze opposte. L'idea di fondo era quella di vedere il ricorso a controfattuali come il tentativo di ristrutturare il contesto delle presupposizioni, reso incoerente dalla supposizione controfattuale, nel modo più economico possibile: il che significa, in termini di semantica modale, ipotizzare che l'antecedente del condizionale sia vero nel mondo «il più possibile simile» al nostro. Asserire un controfattuale significherebbe dunque asserire che il suo conseguente è vero nel mondo il più possibile simile al nostro in cui è vero l'antecedente. Su questa base David Lewis costruiva negli anni settanta una complessa teoria dei controfattuali in cui la nozione di somiglianza fra mondi possibili, assunta come primitiva, portava a un inevitabile realismo circa questi ultimi: essi esistevano per Lewis come il mondo attuale, pur non essendo, per l'appunto, «attuali».

Le ricerche sulle modalità fisiche ricevevano un contributo originale anche da Aldo Bressan, cui si deve un importante volume del 1972 che è un'avanzata trattazione della logica modale quantificata condotta nel solco della logica intensionale carnapiana, applicata a problemi di fondazione della fisica. Applicazioni della semantica di Kripke alla fisica si sono avute anche in altre direzioni. Ciò non stupisce se, con Bas van Fraassen, si osserva che la nozione di mondo possibile, a dispetto della sua apparente vaghezza e «metafisicità», si presta in termini intuitivi a delle naturali applicazioni «empiriche», potendo ricevere interpretazioni quali «situazione fisica», «stato fisico» e così via. Vogliamo qui soffermarci su alcune applicazioni ai fondamenti della logica quantistica sull'origine delle quali si veda II, 7.2.

Alla formulazione di una logica quantistica minimale (**MQL**) si giunge sfruttando un collegamento fra formule e particolari algebre che si presentano nella meccanica quantistica. Intuitivamente, ogni formula esprime un enunciato e nell'interpretazione quantistica ogni enunciato sperimentale ha come intensione un insieme di stati fisici. La logica quantistica, come logica degli enunciati sperimentali, riflette la struttura algebrica degli insiemi degli stati fisici. Matematicamente questi sono sottospazi di uno spazio di Hilbert e su di essi è possibile definire tre operazioni che hanno un significato fisico: l'intersezione $\sqcap$; la formazione, dati due sottospazi x e y, del sottospazio $x \sqcup y$ da essi generato; l'ortocomplementazione, che ad ogni x associa lo spazio $x^{\perp}$ degli elementi ortogonali a quelli di x. Si ottiene in questo modo una struttura

$\mathfrak{Q} = \langle Q, \sqcap, \sqcup, ^{\perp}, 0, 1 \rangle$ che costituisce un reticolo *non distributivo*, dotato di massimo e di minimo e (orto)complementato. La non distributività riflette l'impossibilità, dal punto di vista fisico, di compiere determinate misure. Non mette conto di dare qui esplicitamente gli assiomi e le regole di **MQL**; ci basti osservare che una formula $\mathscr{A}$ è un teorema di **MQL** se e solo se assume il valore 1 in ogni struttura come $\mathfrak{Q}$; questo naturalmente considerando il linguaggio con i soli connettivi &, u, ~, e interpretando & su $\sqcap$, u su $\sqcup$ e ~ su $^{\perp}$. In generale, sono questi i connettivi cui riusciamo a dare esplicitamente un significato fisico. Già a questo punto possiamo concludere che **MQL**, **CPI** o **CPC** sono inconfrontabili, in quanto mentre nella prima è teorema $\mathscr{A}$ u ~ $\mathscr{A}$ ciò non si verifica in **CPI**; e mentre in quest'ultima vale la legge $\mathscr{A} \wedge (\mathscr{B} \vee \mathscr{C}) \equiv (\mathscr{A} \wedge \mathscr{B}) \vee (\mathscr{A} \wedge \mathscr{C})$ che esprime la distributività, questa non vale in **MQL**. In ogni caso entrambe queste logiche sono essenzialmente più deboli di **CPC**.

Nel 1972 il sovietico Dishkant suggeriva per **MQL** un nuovo tipo di realizzazioni da lui dette *modelli semantici* (noi le chiameremo anche, per comodità, realizzazioni alla Dishkant o D-realizzazioni) per le quali dimostrava un teorema di completezza facendo vedere che ogni formula $\mathscr{A}$ di **MQL** è valida rispetto alle D-realizzazioni se e solo se lo è rispetto alle realizzazioni algebriche sopra descritte e che abbiamo detto caratterizzare le **MQL**-tautologie. Per definire le sue realizzazioni Dishkant si serviva di semantiche alla Kripke, ponendo precisamente che un *modello semantico* è una struttura $\mathfrak{G} = \langle\langle I, R, \rangle, G\rangle$ dove $\langle I,R \rangle$ è una struttura-modello nel senso di Kripke in cui R è riflessiva e debolmente transitiva nella quale cioè vale

$$\forall gh(gRh \supset \exists i(hRi \wedge \forall j(iRj \supset gRj))) \qquad g, h, i, j, \in I$$

e $\mathfrak{G}$ associa ad ogni $i \in I$ una valutazione classica soddisfacendo una serie di condizioni che non interessa qui riportare. I è inteso intuitivamente come un insieme di possibili stati di conoscenza, un suo elemento come una collezione di fatti fisici noti in un particolare momento, la relazione R come il possibile scorrere del tempo: il passaggio da uno stato g a uno stato h è connesso con l'esecuzione di un esperimento, sicché gRh verrà inteso come: «se conosciamo ora g, è possibile che più tardi, una volta completato un esperimento, conosceremo h».

Al di là dell'intrinseco interesse della proposta di una «nuova»

semantica per **MQL**, da un punto di vista epistemologico più aderente, per così dire, all'andamento intuitivo dei fatti fisici, mette conto rilevare che sulla base delle proprietà della relazione R Dishkant stabiliva un interessante confronto con la logica intuizionista. La transitività debole della relazione R rende talora impossibile una transizione poniamo da g a i, gRi, anche nel caso si abbia, per un qualche h, tanto gRh quanto hRi: intuitivamente «lo stato di conoscenza non può essere cambiato troppo rapidamente». Nei modelli di Kripke per **CPI**, invece, R è transitiva, sicché il passaggio precedente è sempre possibile: «la conoscenza ottenuta con molti passaggi può essere raggiunta con un passaggio solo... non esistono restrizioni alla quantità di conoscenza che può essere acquisita in un sol colpo».

Leggiamo ora, con Dishkant, $\mathfrak{G} \models_h \mathscr{A}$ come: «dalla collezione di fatti h possiamo dedurre la verità di $\mathscr{A}$» (h «forza» $\mathscr{A}$). Dishkant dimostra un lemma che estende ad *ogni* formula una delle condizioni per $\models$ data solo per formule atomiche: precisamente se $\mathscr{A}$ è una formula qualunque di **MQL**, per ogni $g \in I$ si ha

(1) $$\mathfrak{G} \models_g \mathscr{A} \text{ se e solo se } \forall h(gRh \supset \exists i(hRi \wedge \mathfrak{G} \models_i \mathscr{A}))$$

che intuitivamente si può leggere come affermante che la verità è qualcosa che non può essere cancellata senza essere ripristinata come risultato di una singola osservazione, vale a dire sono possibili solo variazioni ripristinabili degli stati di conoscenza. Viceversa, per un modello intuizionista invece della (1) abbiamo

(2) $$\mathfrak{G} \models \mathscr{A} \text{ se e solo se } \forall h(gRh \supset \mathfrak{G} \models_h \mathscr{A})$$

vale a dire fatti nuovi non possono rimpiazzare i vecchi, i fatti si possono solo accumulare o, altrimenti detto, la verità si conserva indipendentemente dal sorgere di fatti nuovi: «la verità è eterna». In conclusione, secondo l'autore, «la logica quantistica è una logica di fatti ripristinabili che variano lentamente; la logica intuizionista è una logica dell'accumulazione dei fatti».

Sulla base di semantiche alla Kripke come quella di Dishkant è stato possibile affrontare alcuni problemi specifici della logica quantistica, in particolare quello della implicazione. Come detto sopra, dietro alla logica quantistica sta una struttura di reticolo non distributivo e come si sa per strutture di questo tipo non è possibile definire operazioni, come lo pseudocomplemento relati-

vo presente nelle algebre di Heyting, che simulino le proprietà minimali di un buon concetto di condizionale. Numerosi lavori sono stati quindi dedicati al problema di come ovviare a questa impossibilità. Nel 1973 Maria Luisa Dalla Chiara sviluppava un tentativo in questa direzione basandosi su una traduzione di **MQL** in un sistema modale **CPC*** che è un ampliamento di **CPC** ottenuto aggiungendo al suo linguaggio i simboli L e M per gli operatori modali di necessità e possibilità rispettivamente. La traduzione avviene mediante la funzione ρ così definita:

$$\begin{aligned} \rho(A_i) &= \mathrm{LM}A_i \text{ (per ogni lettera proposizionale } A_i \text{ di } \mathbf{MQL}) \\ \rho(\sim \mathscr{A}) &= \mathrm{L} \neg\, \rho(\mathscr{A}) \\ \rho(\mathscr{A} \,\&\, \mathscr{B}) &= \rho(\mathscr{A}) \wedge \rho(\mathscr{B}) \\ \rho(\mathscr{A} \,\mathrm{u}\, \mathscr{B}) &= \mathrm{LM}(\rho(\mathscr{A}) \vee \rho(\mathscr{B})) \end{aligned}$$

che porta la negazione quantistica nella necessità della negazione classica e la disgiunzione quantistica nella necessità della possibilità della disgiunzione classica. Anche **CPC*** risulta semanticamente caratterizzato dalle strutture di Kripke la cui relazione di accessibilità sia riflessiva e debolmente transitiva, sicché valendosi del risultato di Dishkant l'autrice può dimostrare che

$$(*) \qquad \models_{\mathbf{MQL}} \mathscr{A} \quad \text{se e solo se} \quad \models_{\mathbf{CPC^*}} \rho(\mathscr{A})$$

(ossia una formula di **MQL** è valida per **MQL**, se e solo se la sua traduzione è valida per **CPC***).

Considerato che **MQL** può essere estesa in modo standard a un calcolo del primo ordine, se indichiamo con $\bigwedge$ e $\bigvee$ i quantificatori quantistici, poniamo[11]

$$\begin{aligned} \rho(\bigwedge x\mathscr{A}) &= \forall x \rho(\mathscr{A}) \\ \rho(\bigvee x\mathscr{A}) &= \mathrm{LM}\exists x \rho(\mathscr{A}). \end{aligned}$$

[11] In questo caso le realizzazioni algebriche sono reticoli ortocomplementati *completi*, con due operazioni infinitarie $\bigcap$ e $\bigcup$, e la valutazione v viene così estesa:

$$\begin{aligned} \bar{v}\,(\bigwedge x\, \mathscr{A}) &= \bigcap \{v(\mathscr{A}(x/d))\}_{d \in D} \\ \bar{v}\,(\bigvee x\, \mathscr{A}) &= \bigcup \{v(\mathscr{A}(x/d))\}_{d \in D} \end{aligned}$$

dove D rappresenta il dominio della realizzazione e d è un nome per l'individuo d.

L'autrice può a questo punto trarre una prima conclusione: «La logica quantistica è senza alcun dubbio una "logica reale". Anche se storicamente suggerita dalla meccanica quantistica, la sua interpretazione modale mostra chiaramente che essa *possiede un significato logico* che è del tutto indipendente dalla meccanica quantistica.» È a riprova di questa conclusione sullo *status* epistemologico della *quantum logic*, che l'autrice presenta una sua soluzione al problema del condizionale per la logica quantistica e mostra come la proposta traduzione modale permetta di dare una risposta alla questione.

La Dalla Chiara osserva che nei tipi di strutture corrispondenti a **CPC**, **CPI**, **MQL** può essere definita un'operazione $\multimap$ ponendo

$$x \multimap y = \begin{cases} 1 & \text{se } x \leqslant y \\ 0 & \text{altrimenti} \end{cases}$$

che si comporta come una «implicazione stretta» nel senso che ponendo $\bar{v}(\mathscr{A} \supset \mathscr{B}) = \bar{v}(\mathscr{A}) \multimap \bar{v}(\mathscr{B})$ si «evitano» i paradossi dell'implicazione materiale. Ampliamo ora il linguaggio di **MQL** aggiungendo il nuovo connettivo binario $\rightarrow$ e poniamo

$$\bar{v}(\mathscr{A} \rightarrow \mathscr{B}) = \bar{v}(\mathscr{A}) \multimap \bar{v}(\mathscr{B});$$

estendiamo contemporaneamente la traduzione da **MQL** a **CPC*** con l'ulteriore clausola

$$\rho(\mathscr{A} \rightarrow \mathscr{B}) = \rho(\mathscr{A}) \multimap \rho(\mathscr{B}).$$

Il risultato (*) si estende allora in modo diretto anche a questo caso, si ottiene cioè

$$\models_{\mathbf{MQL}} \mathscr{A} \rightarrow \mathscr{B} \quad \text{se e solo se} \quad \models_{\mathbf{CPC^*}} \rho(\mathscr{A}) \multimap \rho(\mathscr{B}).$$

Ne discende che questo tipo di implicazione stretta può essere un candidato al ruolo di implicazione quantistica, dal momento che «essa è *del tutto omogenea* con l'interpretazione modale degli altri connettivi quantistici».

Altre traduzioni dei sistemi di logica quantistica in sistemi modali classici venivano fornite nel 1974 da Goldblatt ed in generale numerose sono state le indagini sia dal punto di vista algebrico

(secondo la prospettiva originale di Birkhoff e von Neumann) sia dal punto di vista semantico più generale sulle logiche quantistiche. Tra di esse ricordiamo in particolare quelle di Simon Kochen ed Ernst Specker della metà degli anni sessanta dedicate alla costruzione di una semantica per le logiche quantistiche con operazioni parziali e all'analisi del problema delle variabili nascoste.

Un ruolo importante – anche se in un ambito di ricerca più modesto – la semantica a mondi possibili ha avuto anche nel campo della logica deontica, che tanto deve alle ricerche di Georg H. von Wright, uno dei padri della moderna logica modale. La logica deontica è in effetti una logica modale priva dell'assioma $\mathscr{A} \supset \mathrm{M}\mathscr{A}$ (che verrebbe a significare «tutto è permesso») e abitualmente dotata in un suo luogo dell'assioma $\mathrm{L}\mathscr{A} \supset \mathrm{M}\mathscr{A}$ («ciò che è obbligatorio è permesso»). I mondi possibili dei modelli deontici sono spesso interpretati come «i migliori tra i mondi possibili» con il risultato di dare un sapore leibniziano anche alle attuali discussioni di filosofia morale. Con questi strumenti è stato possibile affrontare problemi come quello dei rapporti tra enunciati descrittivi ed enunciati normativi e la questione degli obblighi condizionali, quegli obblighi cioè che dipendono dal realizzarsi fattuale di determinate circostanze. È stato proprio per evitare i numerosissimi paradossi che questi tipi di obblighi comportano che ai sistemi originari del tipo presentato da von Wright a partire dal 1951, sono stati sostituiti sistemi di logica deontica diadica in cui l'operatore modale monadico L di obbligatorietà assoluta viene rimpiazzato da un operatore diadico di obbligo condizionale. Connesse a queste ricerche sono tanto le indagini sulla logica degli imperativi, la cui origine risale a Ernst Mally nel 1926 e che furono poi portate avanti ad esempio da Karl Menger nel 1939 e sempre nello stesso anno da Albert Hofstadter e McKinsey, quanto le più recenti ricerche dedicate agli aspetti logici della teoria delle decisioni razionali e dell'azione sviluppate da Vaughan R. Pratt e Segerberg che hanno contatti con la stessa informatica teorica.

Un discorso simile si potrebbe fare anche per altre logiche, quali quella *epistemica* (in cui si introducono operatori del tipo K*a* «*a* sa che...», B*a* «*a* crede che...») quella *proairetica* (la logica della preferenza) quella *erotetica* (la logica degli interrogativi) ed altre, che nel periodo d'anni di cui ci stiamo occupando sono andate via via prendendo forma soprattutto per opera di Hintikka, von

Wright, Prior, Åqvist, ecc. La maggior parte di queste ricerche non tardarono a confluire in un settore profondamente coinvolto nell'impiego della semantica intensionale, quello dell'analisi del linguaggio naturale quale è stata inaugurata da Montague nel 1968 e nel 1970 estendendo la semantica alla Tarski ai linguaggi naturali, mossa questa della cui opportunità (sia detto per inciso) il primo a dubitare era stato proprio Tarski stesso.

Un esempio particolarmente interessante di queste ricerche riguarda l'analisi dei *tempi verbali* che ha dato origine per opera di Prior alla cosiddetta *tense logic*. Fondamentale in questo campo è la lunga analisi condotta nel 1971 da Hans Kamp sulla logica dell'avverbio «ora», e profondamente innovatrice la serie di ricerche di Åqvist e Franz Günther sugli aspetti logici delle nozioni di «cominciare a», «cominciare con», «finire con», «diventare sempre più vero che», ecc. Altro tema di notevole interesse e già affrontato a suo tempo da Russell è quello degli *atteggiamenti proposizionali* (propositional attitudes) cui Hintikka ha dedicato significative ricerche negli anni settanta, sfruttando la sua semantica per la logica modale. E l'elenco potrebbe continuare.

Abbiamo parlato dei tentativi di Scott di recuperare anche le logiche polivalenti nel contesto di un'analisi di tipo intensionale. Il discorso si inquadrava in un ripensamento generale della formalizzazione della logica iniziato da Scott nel 1973, in un lavoro appunto dal titolo *Background to formalization* (*Sfondo della formalizzazione*) in cui venivano riaffrontate in chiave critica le stesse assunzioni di base della formalizzazione: ad esempio l'ipotesi che i linguaggi siano insiemi, che la chiusura rispetto ai connettivi sia illimitata, che le valutazioni siano bivalenti, che le dimostrazioni di completezza giustifichino le semantiche ecc. In questa ottica venivano ripresi il problema della forma delle regole logiche (di qui il collegamento con l'*entailment* cui sopra si è accennato) e le proprietà generali del concetto di valutazione e di conseguenza sviluppate poi sistematicamente da Gabbay. L'obiettivo di Scott era di «misurare» – fornendo alternative concrete – il peso di tutte queste assunzioni dal punto di vista filosofico e matematico.

Tra le assunzioni che Scott poneva in discussione ve ne erano due che riguardano i presupposti esistenziali e il ruolo dei quantificatori; esse sono 1) che tutti i termini hanno una denotazione e 2) che i quantificatori fissano il dominio dell'interpretazione, as-

sunzione quest'ultima (esistere vuol dire essere valore di una variabile quantificata) che sin dagli anni quaranta Quine aveva posto come criterio delle compromissioni ontologiche che si fanno costituendo una teoria. Queste assunzioni erano state da tempo messe in discussione e costituirono il punto di partenza della logica inclusiva e della logica libera, di cui ora ci occuperemo.

La semantica della logica standard contiene due condizioni che si traducono immediatamente in profondi impegni di carattere ontologico. La prima richiede che il dominio di ogni realizzazione *non sia vuoto*: su questa base possiamo facilmente dimostrare la validità della formula $\exists x(x = x)$ e quindi giustificare il fatto che l'esistenza di almeno un oggetto *sia una verità logica*. La seconda richiede invece che la funzione di interpretazione assegni un oggetto del dominio a tutte le costanti individuali (o, se il calcolo è privo di costanti, a tutte le variabili individuali): si può così affermare che la logica standard esclude la possibilità di termini singolari non denotanti, come le descrizioni del tipo «l'attuale re di Francia», ecc. Su una parafrasi di espressioni di questo tipo che esplicitamente negava loro valore esistenziale era fondata – come si ricorderà – l'analisi di Russell delle descrizioni.

I problemi cui queste condizioni danno luogo sono di duplice natura. Da un lato infatti si può sollevare la questione dell'adeguatezza del calcolo come strumento di indagine della logica del discorso ordinario, in vista del fatto che tale discorso contiene spesso riferimenti a oggetti «fittizi»; mentre dall'altro, in sede più propriamente filosofica, ci si può chiedere se sia giustificata questa stretta connessione fra logica e ontologia. L'esistenza di questi due ordini di problemi ha portato spesso gli autori interessati a fornire risposte diverse, e talvolta difficilmente conciliabili.

Intorno al 1950 l'interesse si accende sulla prima condizione di cui sopra: in una serie di articoli di Mostowski, Church, Theodore Hailperin e Quine, si formulano varie proposte relative alla costruzione di un sistema di teoria della quantificazione i cui teoremi risultino validi anche nel dominio vuoto. Tutte queste proposte non superano tuttavia una difficoltà relativa alla costruzione di una procedura che permetta di valutare in modo soddisfacente le formule aperte nel dominio vuoto, difficoltà la cui natura è più riposta di quanto non sembri. Si osservi infatti che nella semantica standard la verità è concepita in modo corrispondentista (per in-

tenderci, alla Tarski): la formula $P\bar{x}$ è vera in una determinata interpretazione (per una particolare valutazione delle variabili individuali) se e solo se l'oggetto assegnato a x è contenuto nella classe dei P, ossia, in termini più intuitivi, se *si verifica che* Px (Px è un fatto) in quella realizzazione e per quella valutazione. Ora però se ad x non è assegnato alcun valore, come è possibile decidere se Px è vera (cioè se corrisponde a un fatto) sulla base di una simile nozione corrispondentistica di verità (o se si preferisce di una simile nozione realista di «fatto»)? È evidente dunque che l'ammissione di termini singolari non denotanti comporta quanto meno una rimeditazione complessiva del concetto di interpretazione, che solo così potrà essere risolto nel caso particolare del dominio vuoto, dove tutti i termini singolari sono necessariamente non denotanti.

Negli anni cinquanta i tempi tuttavia non erano maturi per una impostazione così generale del problema che fu invece affrontato in termini assai riduttivi: da un lato, proponendo (Church, Hailperin, Quine) la costruzione di sistemi di quantificazione che ammettessero solo teoremi *chiusi* (cioè senza variabili libere), dall'altro sostenendo (Mostowski, e successivamente Hubert Schneider) che le formule aperte sono tutte vere nel dominio vuoto, per semplice estensione (peraltro non altrimenti motivata) a questo dominio del metateorema di chiusura. I sistemi così costruiti sono noti comunemente, con espressione coniata da Quine, come sistemi *inclusivi* (s'intende, del dominio vuoto).

Nel 1956 un articolo di Henry Siggins Leonard vivamente critico nei confronti della regimentazione russelliana e quineana, comincia a delineare le caratteristiche essenziali di una nuova logica dei termini singolari, in cui veniva a cadere l'ipotesi che tali termini siano tutti denotanti. Tale logica prende via via una forma sempre più precisa attraverso contributi di Hintikka, Hailperin, Hugues Leblanc e Karel Lambert e riceve da quest'ultimo la denominazione di «logica libera (da assunzioni esistenziali)». La caratteristica fondamentale dei sistemi di logica libera consiste nel loro rifiuto della cosiddetta *legge di specificazione* (e della corrispondente *regola di esemplificazione universale*) $\forall x \mathscr{A} \supset \mathscr{A}(x/y)$, che viene sostituita con la legge più debole $(\forall x \mathscr{A} \wedge E!y) \supset \mathscr{A}(x/y)$, dove $E!$ è un predicato monadico che va letto «esiste» e che in una logica dell'identità può essere definito nel modo seguente: $E!y =_{Df} \exists x (x = y)$. La legge di specificazione debole ci permette di passare ad

esempio da «Tutti gli *x* sono *P*» a «*a* è *P*» (o, il che è lo stesso, da «*a* è *P*» a «Esiste un *x* che è *P*») solo se *a* esiste, e quindi blocca inferenze illecite come quella da «Pegaso è Pegaso» a «Esiste qualcosa che è Pegaso».

Come abbiamo detto però il vero problema connesso ai termini singolari non denotanti è di natura semantica e riguarda in definitiva i rapporti tra verità (in particolare, verità logica) ed esistenza. Più che il calcolo in sé è pertanto opportuno discutere le varie interpretazioni che di esso sono state date e che possono essere raccolte in quattro filoni principali.

Il primo (e anche il meno seguito) è rappresentato sostanzialmente da Rolf Schock, che in numerosi articoli pubblicati dal 1961 in avanti venne proponendo e articolando la posizione seguente (in cui si riconosceranno evidenti influssi russelliani): una formula atomica è vera in una determinata realizzazione (e rispetto a una data valutazione) se e solo se corrisponde a un fatto che si verifica in quella realizzazione (e rispetto a quella valutazione) e dunque se una tale formula contiene termini singolari non denotanti, e di conseguenza non può corrispondere a un fatto, essa è *falsa*. Evidentemente, sulla base di una tale valutazione delle formule atomiche, sarà possibile valutare anche tutte le formule complesse contenenti termini singolari non denotanti; si tenga conto comunque che per quanto riguarda la valutazione di formule quantificate Schock considera rilevanti solo gli esempi di tali formule costruiti con termini singolari *denotanti* (ovvero, il che è lo stesso, solo le realizzazioni alternative che assegnano un valore alla variabile quantificata) ovviamente in vista della *portata esistenziale* che egli (d'accordo in questo con la semantica standard) intende dare ai quantificatori (tale caratteristica sarà conservata anche in tutte le semantiche alternative).

Lo strumento fondamentale del secondo filone interpretativo è costituito dalla nozione di «dominio esterno», ossia di un dominio di oggetti inesistenti che possono essere assegnati come *denotata* ai termini singolari non denotanti. Sulla base di questa nozione è molto semplice costruire una semantica libera: basterà infatti conformarsi al modello di una semantica standard in cui però i quantificatori siano «ristretti» a un sottoinsieme del dominio di interpretazione. Sarà molto semplice in particolare spiegare *perché* una formula come $x = x$ è vera anche quando x non esiste: infatti, che

x esista o meno (ossia che il valore assegnato a x appartenga al dominio interno o a quello esterno) tale valore coinciderà ovviamente con se stesso. Ciononostante, quello che otterremo in tal modo potrà difficilmente essere considerato soddisfacente da un punto di vista filosofico (e in tal senso si sono espressi anche due fra i principali artefici della semantica dei domini esterni, Leblanc e Thomason) in quanto per questa via possiamo al più evitare la questione fondamentale delineata sopra, non certo risolverla. Se il problema è «Come possiamo valutare degli enunciati che contengono termini singolari non denotanti, dal momento che essi non possono corrispondere ad alcun stato di cose?», è evidente che una risposta del tipo «Possiamo valutare tali enunciati perché essi corrispondono (o non corrispondono) a degli stati di cose relativi agli oggetti *fittizi* cui si riferiscono», mentre provoca uno stravolgimento delle nozioni stesse di «stato di cose» e di «termine non denotante», vanifica addirittura la domanda e la difficoltà che essa tentava di esprimere. È opportuno comunque ricordare che esistono, oltre a quella di Leblanc e Thomason, due varianti di questo filone semantico: quella di Nino B. Cocchiarella (già presente in germe in alcune osservazioni di Church e di cui abbiamo già parlato a proposito delle logiche temporali) in cui non abbiamo due domini disgiunti, ma piuttosto un unico dominio (cui appartengono tutti gli oggetti *possibili*) del quale la collezione degli esistenti è un sottoinsieme (in questo modo le analogie con la teoria «ristretta» della quantificazione sono ancora più evidenti); e quella di Scott, che in un certo senso realizza un compromesso con la «teoria dell'oggetto prescelto» di Frege-Carnap, in quanto costruisce dei domini esterni che contengono esattamente un oggetto, il quale funge da *denotatum* improprio di *tutti* i termini singolari non denotanti.

Assai più stimolante ai fini della ricerca di una soluzione soddisfacente alla nostra questione fondamentale si presentava la linea impostata da van Fraassen, in una serie di articoli pubblicati fra il 1966 e il 1969. Per van Fraassen una formula $\mathscr{A}$ contenente termini singolari non denotanti non può ricevere *in prima istanza* alcun valore di verità. Ciononostante, essa può ricevere un valore di verità come risultato di una sorta di esperimento mentale: possiamo cioè supporre di assegnare a tutte le componenti atomiche di $\mathscr{A}$ che in una data realizzazione risultano invalutabili (perché ap-

punto contengono termini singolari non denotanti) un valore di verità *qualsiasi*, e quindi assegnare un valore di verità ad $\mathscr{A}$ nella realizzazione data se e solo se per ogni valutazione delle componenti atomiche di $\mathscr{A}$ che coincida con la valutazione associata alla realizzazione stessa *là dove questa è definita* e che la completi nel modo sopra specificato, $\mathscr{A}$ riceve sempre lo stesso valore di verità.

Il discorso di van Fraassen può essere interpretato mediante una metafora assai suggestiva: possiamo pensare cioè ai termini singolari non denotanti semplicemente come a dei termini sui quali non abbiamo alcuna informazione. Di essi dunque non possiamo *in prima istanza* dire nulla; può darsi però che, per ogni completamento possibile della nostra informazione che ci ponga in grado di dire qualcosa relativamente a tali termini, certi enunciati che li contengono risultino invariabilmente veri o invariabilmente falsi. In tal caso, in base alla posizione di van Fraassen, siamo autorizzati a considerarli veri o falsi *già adesso*.

La proposta è indubbiamente interessante, ma conviene mettere subito in rilievo quello che costituisce un suo limite. Lo strumento che permette di valutare le formule logicamente (e non fattualmente) vere o false è il concetto di *supervalutazione* che fu utilizzato da van Fraassen in numerosi altri contesti, che agisce di fatto solo a livello di enunciati non analizzati. Una supervalutazione infatti è costruita su valutazioni che assegnano un valore di verità qualsiasi alle componenti atomiche invalutabili di una formula, *indipendentemente dalla loro struttura*. Per questo motivo, le uniche formule logicamente vere o false che una supervalutazione riesce a valutare *in modo naturale* sono tautologie e contraddizioni; là dove invece la *struttura* delle formule atomiche ha una rilevanza essenziale la supervalutazione non può arrivare. Un esempio molto semplice di questa limitazione si ha considerando la formula $x = x$. Se in una data realizzazione x è un termine singolare non denotante, certamente la supervalutazione – per come essa è stata definita – non potrà assegnare il valore Vero a tale formula, dal momento che essa è atomica e quindi, nelle valutazioni ideali che completano quella associata alla realizzazione data, riceverà ora il valore Vero, ora il valore Falso. Per ovviare a questo inconveniente, van Fraassen fu costretto a ricorrere a un espediente *ad hoc*, ammettere cioè che tutte le valutazioni ideali di cui sopra *per definizione* assegnassero il valore Vero alla formula in questione, cosicché

la supervalutazione costruita su di esse le assegnasse ancora il valore Vero (analoghe condizioni dovevano essere introdotte per validare gli altri assiomi relativi alla quantificazione e all'identità). In questo modo però la grande plausibilità di questa linea di soluzione veniva ad essere in gran parte perduta: $x = x$ infatti non è vera (quando x non esiste) in base al successo di un esperimento mentale, ma semplicemente in base a una definizione.

Il quarto e ultimo filone interpretativo, rappresentato in particolare da Leblanc, raccoglie e in certo senso porta alle estreme conseguenze i suggerimenti di van Fraassen, giungendo alla costruzione della cosiddetta *semantica vero-funzionale*, nella quale i domini sono completamente eliminati in favore di valutazioni arbitrarie delle formule atomiche del linguaggio. L'idea che le formule logicamente valide o contraddittorie siano quelle invarianti rispetto a tutte le trasformazioni possibili dei valori di verità delle loro componenti atomiche, applicata fino in fondo, permette di eliminare qualsiasi riferimento al piano della «realtà»; ma si ripresentano allora puntualmente le difficoltà che già van Fraassen aveva dovuto affrontare relativamente alla specificità di certe strutture enunciative (quelle per così dire intrinsecamente quantificazionali) e le stesse soluzioni *ad hoc*.

La logica libera ha trovato diverse applicazioni tra le quali ricordiamo la costruzione di teorie delle descrizioni più soddisfacenti dal punto di vista intuitivo di quelle tradizionali;[12] la costruzione di logiche modali predicative; i tentativi di applicazione alla teoria degli insiemi. In quest'ultimo campo, oltre a un articolo di Scott in cui si fornisce un'interpretazione «libera» della teoria delle classi virtuali di Quine, sono da ricordare interessanti lavori di R.H. Thomason, Ermanno Bencivenga e A. Trew, per limitarci ad alcuni dei più importanti. Va infine ricordato che nel contesto della logica categoriale la problematica della logica libera ha avuto oggi sviluppi estremamente interessanti soprattutto ad opera di André Joyal che ha posto in luce collegamenti con la teoria dei fasci.

[12] Nel senso che in esse, ad esempio, si può affermare, a differenza che in quella di Russell, la verità di «l'x tale che Px = l'x tale che Px» anche se un tale x non esiste o non è unico.

6. LA TEORIA DEGLI INSIEMI DA GÖDEL A COHEN

Malgrado la centralità di cui indubbiamente gode la teoria degli insiemi dal punto di vista fondazionale, non si può dire che i risultati ottenuti da Gödel sulla coerenza di *AS* e *IGC* alla fine degli anni trenta avessero aperto una stagione di intense ricerche di carattere matematico e metamatematico in questo campo. Come si rivela ad un esame anche superficiale della letteratura, essi furono seguiti invece da un periodo di sostanziale stagnazione che durò per circa vent'anni. Indubbiamente cause esterne non furono estranee a questo fatto anche se un'analoga situazione di rallentamento non si verificò in discipline collaterali. Le ragioni più profonde ci sembra vadano cercate altrove e per chiarirlo conviene forse fare un passo indietro nella nostra storia.

Abbiamo descritto a suo tempo lo sviluppo dell'assiomatizzazione della teoria degli insiemi e della sua formalizzazione, con il costituirsi della teoria $\mathfrak{ZF}$ e parallelamente della teoria delle classi $\mathfrak{NBG}$. Come avremo occasione di vedere in dettaglio più avanti, questi non furono gli unici tentativi di assiomatizzare la teoria degli insiemi e delle classi, ma innegabilmente si presentarono come i più influenti sugli sviluppi successivi. Non si deve però credere che dall'inizio degli anni trenta in poi fosse un fatto riconosciuto che

1) tutta intera la matematica fosse ricostruibile all'interno dell'universo insiemistico e quindi che – in questo senso – la teoria degli insiemi costituisse la *fondazione* della matematica né che

2) il modo più diretto e vicino alla pratica del matematico per descrivere l'universo insiemistico fosse proprio quello offerto da $\mathfrak{ZF}$ o $\mathfrak{NBG}$.

Se prima Frege e poi soprattutto i *Principia Mathematica*, avevano mostrato come l'intero corpo della matematica classica fosse riproducibile in termini insiemistici, nel senso che ogni oggetto di cui il matematico classico si occupa aveva un suo «modello» (o se si preferisce una sua «copia») all'interno dell'universo degli insiemi, diverso era il problema per quanto riguardava la matematica che *si andava facendo* e l'opportunità di vedere nella teoria degli insiemi il punto di riferimento obbligato sul piano operativo.

Come sappiamo, la svolta decisiva della matematica a partire dai primi decenni del secolo era stata essenzialmente una svolta

astratta, che comportava la generalizzazione di nozioni e metodi al di là dei contesti in cui si disponeva di una sorta di riscontro «empirico», attraverso l'introduzione di nuovi oggetti e di nuove teorie. Si pensi al sorgere dell'Analisi funzionale, della topologia generale, dell'algebra astratta, ecc.: tutti campi in cui nozioni incontrate in casi concreti venivano analizzate in una prospettiva più astratta che permetteva di modellare matematicamente situazioni più generali. In questo contesto la prospettiva assiomatica di Hilbert non aveva tardato a rivelarsi uno strumento essenziale, in quanto offriva la possibilità di studiare questi nuovi oggetti che non avevano riscontro concreto *diretto*, in un modo che rendeva pienamente la loro generalità; anche nel caso degli insiemi il metodo assiomatico ebbe un ruolo fondamentale e non stupisce quindi che nozioni centrali come quelle di spazio metrico, di spazio topologico, di gruppo, di anello, di campo, ecc., trovassero anch'esse la loro prima presentazione in termini assiomatici.

Questo però non era sufficiente: in primo luogo in quanto dal punto di vista matematico è estremamente innaturale non poter considerare gli enti di cui parlano le teorie matematiche come oggetti, studiandone i mutui rapporti; in secondo luogo in quanto la stessa nozione di teoria assiomatica e i metodi per studiarne coerenza, completezza, e altre proprietà metamatematiche, rimandavano sempre alla nozione di *interpretazione* e di *modello*. L'affermarsi della semantica tarskiana costituiva in un certo senso la consacrazione a livello formale di questo fatto.

Qual era l'universo in cui tutte le teorie potevano interpretarsi e in cui «vivevano» i loro modelli? È come universo generale in cui interpretare ogni teoria matematica che la teoria degli insiemi si presentò nel suo ruolo fondazionale, ed è questo un passaggio che avvenne automaticamente, senza l'esplicitazione di presupposti e di obiettivi di carattere epistemologico generale, diversamente da come era successo, ad esempio, per il programma logicista.

Dopo la sistemazione di Zermelo in cui la stessa nozione intuitiva di universo degli insiemi trovava una traduzione assiomatica, i matematici venivano a disporre di un modo ancora più determinato per concepire questo ruolo fondazionale: quello di vedere ogni specifica teoria matematica come *traducibile* entro quella degli insiemi. Era all'interno di questa teoria che si trattava di trovare

modi per definire i diversi oggetti, a partire dal puro concetto di insieme, fossero questi funzioni, spazi, algebre, ecc.

Il processo fu lungo e solo col tempo si ebbe la prova concreta che la semplice nozione di appartenenza era sufficiente a definire ogni altra nozione fondamentale. Un ruolo centrale, come abbiamo ricordato, ebbe in questo sviluppo la definizione del concetto di coppia ordinata, data da Norbert Wiener e successivamente da Kasimierz Kuratowski, che permetteva la riduzione della teoria delle relazioni e delle funzioni a quella degli insiemi. Ruolo analogo ebbe lo sviluppo della teoria degli ordinali dato da von Neumann nel 1923, mediante il quale gli ordinali venivano identificati con gli insiemi ben fondati e transitivi, e divenivano così *insiemi* e non più classi come nella definizione cantoriana; lo stesso vale per la conseguente definizione dei cardinali come ordinali iniziali. Questo lento processo di traduzione non aveva tardato a porre in luce il ruolo di principi insiemistici controversi, quali l'assioma di scelta *AS* e l'ipotesi del continuo *IC*, nella dimostrazione di teoremi specifici di Analisi, topologia, algebra, mostrando *in concreto* il peso che queste ipotesi avevano nella pratica matematica astratta, al di là delle dispute di principio del primo decennio del secolo.

Esempi significativi per l'ipotesi del continuo vennero presentati già nel 1934 da Wacław Sierpinski nel suo classico *Hypothèse du continu* (*Ipotesi del continuo*), che mostrava in modo drammatico il ruolo che tale ipotesi aveva nello studio di problemi classici riguardanti la teoria delle funzioni. Parallelamente, per quanto riguarda l'assioma di scelta, si andava scoprendo il suo ruolo decisivo nell'algebra (ad esempio, per dimostrare che ogni campo ha un'unica chiusura algebrica, che ogni spazio vettoriale ha una base, che ogni sottogruppo di un gruppo libero è libero, ecc.); nell'Analisi funzionale e nella teoria della misura (per dimostrare il classico teorema di Vitali secondo il quale esistono insiemi limitati di reali che non sono misurabili nel senso di Lebesgue, o per il teorema di Hahn-Banach sulla prolungabilità dei funzionali lineari definiti su spazi vettoriali); nella topologia (ad esempio per il teorema di Thikhonov); o ancora nella teoria dei reticoli e nella logica, ad esempio per dimostrare il teorema del filtro massimale o che ogni teoria coerente si estende ad una completa, come afferma il teorema di Lindenbaum.

Non erano mancati in questo contesto anche gli esempi che mostravano lo scarto tra le situazioni concrete e la loro modellizzazione insiemistica; il più celebre al riguardo (nel contesto del problema generale della misura) era già stato offerto addirittura nel 1924 da Banach e Tarski i quali dimostrarono che *AS* comporta che una sfera nello spazio reale ($\mathbf{R}^3$) può essere decomposta in due sfere ciascuna delle quali congruente con la sfera di partenza.

Tutti questi fatti mostravano chiaramente che, se da una parte la teoria degli insiemi costituiva realmente un *linguaggio universale* in cui ricostruire la matematica, all'interno di essa si presentavano interrogativi che avevano un immediato significato logico: quali erano i rapporti di *AS* e di *IGC* con gli altri assiomi della teoria e entro che limiti essi erano essenziali per ottenere risultati matematici specifici come quelli sopra citati? Più in generale ancora: che tipi di ipotesi, aggiuntive o alternative, si possono fare sull'universo degli insiemi? Sul piano epistemologico si verificava d'altra parte una situazione in qualche modo paradossale per cui molti matematici di sicura professione formalista si comportavano come inveterati platonisti, accettando come contesto della loro pratica matematica una teoria – come quella degli insiemi – nel cui cuore è inscritta un'assunzione di tipo fortemente realista.

Sul versante più pratico si poneva l'altro problema cui sopra accennavamo che è quello relativo al *modo* di descrivere l'universo degli insiemi in maniera tale che sul piano operativo non ci fosse urto tra la ricostruzione insiemistica e l'ordinario procedere matematico. Già abbiamo visto come sul problema dell'esistenza di *Urelemente* si fosse scontrata una concezione più vicina alla pratica, quella di Zermelo, e una più purista, come quella di Fraenkel. Ma l'antitesi più profonda era un'altra: quella tra i sostenitori della teoria degli insiemi come teoria matematica i cui oggetti sono tutti della stessa sorta e tra i quali è definita un'unica relazione, quella di appartenenza, e invece i sostenitori della teoria dei tipi come contesto in cui sviluppare la matematica.

La seconda opzione era chiaramente inevitabile per chi professava un atteggiamento di tipo predicativista ed in questa prospettiva i *Principia Mathematica*, depurati dall'assioma di riducibilità, rappresentavano un punto di riferimento ineliminabile. È da qui che avevano preso le mosse Weyl e successivamente Lorenzen e

tutti coloro che miravano allo sviluppo di una matematica predicativista. Meno naturale sembrerebbe questa scelta per un platonista, o più in generale per un matematico senza scrupoli di carattere predicativista. Questo era stato l'atteggiamento dei matematici tedeschi ai primi anni del secolo, che contrapponevano il matematico Cantor al logico Russell (non nascondendo una cordiale antipatia per quest'ultimo) ma non si può dire altrettanto per altri paesi e per altri tempi. Già abbiamo parlato del ruolo di Chwistek e Ramsey nell'isolare all'interno dei *Principia* la struttura semplice dei tipi, ma va sottolineato che nessuno dei due concretamente sviluppò il sistema che invece fu presentato per la prima volta in modo dettagliato solo nel 1929 da Carnap nel suo *Abriss der Logistik* (*Compendio di logistica*) e a cui si ispirarono successivamente Quine nel 1934 e Church nel 1940.

Se l'atteggiamento di Chwistek e Ramsey era stato essenzialmente motivato da considerazioni filosofiche, altri matematici – e l'esempio più cospicuo per noi è costituito da Turing – erano andati più in là e avevano visto nella teoria semplice dei tipi un linguaggio *concretamente utilizzabile* dai matematici, sostenendo che la teoria russelliana dei tipi «era stata largamente anticipata dall'uomo preistorico» e rappresentava un dato naturale, insito nella struttura del linguaggio e del pensiero. Così negli anni quaranta Turing sviluppò sistematicamente diverse forme di teoria dei tipi, tra l'altro dimostrando in un articolo rimasto inedito un teorema di normalizzazione per il calcolo con tipi. Questi lavori culminarono nel 1948 con la presentazione di un articolo *Practical Forms of Type Theory* (*Forme pratiche della teoria dei tipi*) che mostrava come per Turing non fosse affatto ovvia l'opzione che la maggior parte dei matematici aveva già compiuto per la teoria degli insiemi rispetto al linguaggio dei tipi e questo è interessante non solo perché in tempi recenti – nel contesto della teoria delle categorie e dell'informatica – la precisione e la duttilità di questi linguaggi si è ripresentata in forma nuova, ma anche perché contribuisce a spiegarci come le indagini sulla teoria assiomatica degli insiemi abbiano avuto un andamento singolare, che riflette le incertezze che per molto tempo i matematici avevano avuto nei confronti della teoria.

È su questo sfondo che si situavano i lavori di Gödel sul modello dei costruibili e la coerenza di *AS* e *IGC* di cui abbiamo parlato (v, 2), ed il fatto paradossale è che sulla comunità matematica nel suo

complesso essi ebbero tutto sommato un effetto più negativo che positivo. Conseguenza dei risultati di Gödel era che alcune proposizioni fondamentali della teoria descrittiva degli insiemi (riguardanti quegli oggetti matematici propri della matematica classica tradizionale) risultavano non dimostrabili né decidibili all'interno di $\mathfrak{Z}\mathfrak{F}$ + *AS*. Ciò sembrava confermare l'idea che all'interno della teoria degli insiemi molti problemi importanti fossero indecidibili in vista della eccessiva generalità e che quindi non fosse proficuo lavorare su questi temi. La teoria rimaneva così qualcosa di essenzialmente estraneo al corpo vivo della matematica.

L'atteggiamento di Gödel al riguardo era diverso, ma sostanzialmente isolato all'interno della comunità dei matematici, in quanto si fondava su una convinzione decisamente platonista che lo portava a vedere nei risultati di indipendenza non la prova di una irrimediabile indeterminatezza della teoria, quanto l'esigenza di cercare gli assiomi opportuni per descriverne l'universo. Questo spiega fra l'altro perché, malgrado Gödel avesse ottenuto risultati significativi già negli anni trenta in direzione della indipendenza di *IGC*, egli fosse molto restio a parlarne in pubblico nel timore che si andasse confermando un atteggiamento di sfiducia nei riguardi della teoria. Gli sviluppi più rilevanti nel campo della teoria assiomatica degli insiemi, nel periodo che intercorre fra i lavori di Gödel e quelli di Cohen, riguardano piuttosto aspetti più particolari e di portata meno rivoluzionaria, anche se in parte contribuirono poi alla realizzazione del risultato di Cohen.

Prescindendo dalle ricerche di carattere più dichiaratamente assiomatico, di cui parleremo nel prossimo capitolo, *grosso modo* possiamo dire che in questo periodo l'attenzione venne rivolta essenzialmente all'analisi più ravvicinata dei vari tipi di tecniche per costruire modelli della teoria degli insiemi. Gödel aveva studiato l'universo dei costruibili utilizzando il metodo dei modelli interni, metodo che permetteva un approccio rigorosamente finitista, di tipo sintattico: i costruibili costituivano una classe definita da una ben precisa formula $L(x)$ che risultava assoluta, nel senso che la soddisfacibilità o meno da parte di un individuo della proprietà $L(x)$ non dipendeva dall'universo ambiente in cui ci si muoveva. In questo contesto, come si ricorderà, vedere i costruibili come modello significava studiare la relativizzazione delle formule rispetto alla proprietà $L(x)$. Risultava però interessante studiare

modelli di tipo diverso ed indagare se potevano esistere *insiemi* e non solo *classi* modello di $\mathfrak{ZF}$, avvicinandosi di più alla nozione generale di modello propria della semantica tarskiana. Nella prospettiva della teoria generale dei modelli, un modello per $\mathfrak{ZF}$ è semplicemente una coppia $\mathfrak{M} = < M, E >$ in cui risultano veri gli assiomi di $\mathfrak{ZF}$. Si pone naturalmente la domanda di quale metateoria usare per studiare modelli in questo senso di $\mathfrak{ZF}$. Più in generale, come è possibile utilizzare $\mathfrak{ZF}$, o qualche sua estensione, come metateoria per studiare modelli di teorie degli insiemi?

Con una tecnica analoga alla gödelizzazione è possibile codificare (non solo nel caso della teoria degli insiemi, ma per ogni teoria) il linguaggio in termini insiemistici, associando ad ogni simbolo un insieme e ricostruendo termini e formule come n-uple di insiemi e quindi in ultima istanza come insiemi. A questo punto operazioni e relazioni sintattiche fra formule diventano operazioni e relazioni fra insiemi, cosicché $\mathfrak{ZF}$ si presenta come metateoria sintattica universale (di fatto basta molto meno, come sappiamo da Gödel) ma la cosa importante è che lo stesso può avvenire per la semantica, cosicché se consideriamo strutture che sono insiemi possiamo definire una relazione *insiemistica* $\models$ tale che il fatto che $\mathscr{A}$ sia vera in $\mathfrak{M}$ si traduce nel fatto che la relazione $\models$ vale fra $\mathfrak{M}$ e l'insieme $\ulcorner\mathscr{A}\urcorner$ associato alla formula $\mathscr{A}$. In questo senso $\mathfrak{ZF}$ si presenta come *semantica universale per ogni teoria*, quindi in particolare per se stessa. Con un'avvertenza, però: la definibilità di $\models$ vista come relazione non comporta in nessun modo che si possa trovare una formula insiemistica con una sola variabile libera che risulti una definizione di verità per $\mathfrak{ZF}$, cosicché l'insieme dei «numeri di Gödel» delle formule vere nell'universo degli insiemi non è un insieme definibile in $\mathfrak{ZF}$. Questa non è che l'immediata traduzione del teorema di Tarski per $\mathfrak{ZF}$. Il fatto importante è che per $\mathfrak{NBG}$ non risulta definibile neppure una relazione come la $\models$ e questo è strettamente collegato ad una differenza fra $\mathfrak{ZF}$ e $\mathfrak{NBG}$ che riguarda l'assiomatizzazione: $\mathfrak{NBG}$ è finitamente assiomatizzata mentre, come si ricorderà, $\mathfrak{ZF}$ non lo è.

Questo mostra una differenza profonda fra $\mathfrak{NBG}$ e $\mathfrak{ZF}$, che però non riguarda il contenuto matematico delle due teorie. Come dimostrato infatti da I. Novak e Mostowski nel 1950 riprendendo lavori dell'anno precedente di H. Wang, $\mathfrak{NBG}$ è un'estensione *conservativa* di $\mathfrak{ZF}$ per quanto riguarda gli enunciati che non contengo-

no variabili per classi. Ma torniamo ai modelli delle teorie degli insiemi.

Ricorrendo ancora al teorema di Gödel, è immediato che non potremo mai dimostrare in $\mathfrak{ZF}$ che esiste una struttura $<M, E>$ dove M è un insieme, per cui $<M, E> \models \ulcorner \mathscr{A} \urcorner$ per ogni $\mathscr{A}$ assioma di $\mathfrak{ZF}$. Come dimostrato da Mostowski, ciò è possibile all'interno di una estensione di $\mathfrak{NBG}$, di solito indicata con $\mathfrak{M}$, introdotta da A.P. Morse nei primi anni cinquanta e presentata per la prima volta in appendice al volume *General Topology* (*Topologia generale*) di J. Kelley del 1955. (Solo dieci anni dopo lo stesso Morse, nel volume *A Theory of Sets* [*Una teoria degli insiemi*] darà una formulazione autonoma del proprio sistema.) Il sistema $\mathfrak{M}$ si può considerare una versione *impredicativa* di $\mathfrak{NBG}$. Come si ricorderà infatti il gruppo B degli assiomi di Gödel è equivalente all'assunzione di un assioma di comprensione per formule in cui non compaiono quantificate variabili per classi, ed è questo che dà allo schema la sua valenza predicativa (una classe non è mai definita ricorrendo alla totalità delle classi). Morse lascia cadere la limitazione sulle formule e accetta quindi lo schema nella sua forma impredicativa. Chiaramente $\mathfrak{M}$ è un'estensione *propria* di $\mathfrak{NBG}$ e risulta interessante domandarsi di quanto sia più forte di essa. La risposta è molto semplice: $\mathfrak{M}$ si ottiene da $\mathfrak{NBG}$ aggiungendo l'assioma AI che afferma l'esistenza di almeno un inaccessibile, dove un cardinale κ è (fortemente) *inaccessibile* quando risulta maggiore di ogni 2^λ per $\lambda < \kappa$ e non si può ottenere come somma di meno di κ cardinali tutti minori di κ (condizione quest'ultima che definisce i cosiddetti cardinali *regolari*).

Questo fatto è già una prima indicazione dello stretto intergioco tra possibilità di riflettere le proprietà dell'universo V (quale appunto la proprietà di essere modello) e l'assunzione di grandi cardinali, che costituisce un tema di ricerca che nasce negli anni sessanta e di cui avremo occasione di parlare più avanti. Ma torniamo ora alle strutture che possono risultare modelli di $\mathfrak{ZF}$ o di suoi frammenti.

Dal punto di vista della teoria generale dei modelli, come abbiamo detto, le interpretazioni del linguaggio di $\mathfrak{ZF}$ sono le strutture $<M, E>$ dove E è una relazione binaria definita su M. All'interno di questa classe generale di interpretazioni viene però naturale considerare due sottoclassi. La prima è quella delle strutture *standard*, che sono le strutture $<M, E>$ dove E coincide

con la restrizione a M della relazione di appartenenza $\in$. Ciò significa che l'appartenenza entro il modello è la stessa della metateoria in cui lo studiamo, cosicché il modello stesso è identificato dal solo M. Nel caso delle interpretazioni standard possiamo provare in modo diretto che la nozione di verità coincide con la verità nell'universo della relativizzazione degli enunciati nel senso che per ogni formula $\mathscr{A}$ e ogni $a_1,...,a_n$ in M avremo l'equivalenza $<M, E> \models \mathscr{A}(a_1, ..., a_n)$ se e solo se $\mathfrak{ZF} \models \mathscr{A}^M$, dove $\mathscr{A}^M$ è la formula $\mathscr{A}$ relativizzata all'insieme M.

Intuitivamente sono le interpretazioni standard i possibili modelli cui si pensa interpretando $\mathfrak{ZF}$; risulta quindi estremamente interessante il teorema che Mostowski dimostrò nel 1949 – oggi noto come *teorema di isomorfismo* – che afferma che ogni struttura $<M, E>$, dove E è una relazione ben fondata che soddisfa l'assioma di estensionalità, è isomorfa ad una struttura standard transitiva. M è transitivo quando, dato $a \in M$, si ha anche per ogni $b \in a$ che $b \in M$. Questo significa che ogni elemento di M è una collezione di elementi che appartengono sempre a M. Nel caso di interpretazioni standard è ovvio che la transitività scende immediatamente dal fatto di soddisfare l'assioma di estensionalità, cosicché i possibili candidati ad essere modelli standard di $\mathfrak{ZF}$ sono gli insiemi transitivi.

Particolare interesse fra le strutture transitive hanno i modelli *naturali*, vale a dire gli universi parziali costituiti dai livelli della gerarchia di von Neumann. È questa la seconda sottoclasse cui accennavamo. Questi modelli furono intensamente studiati negli anni cinquanta da Tarski, che li isolò nel 1956, e dalla sua scuola (Montague, Vaught, Scott, ecc.). Si noti che l'esistenza di un modello standard transitivo non comporta in nessun modo l'esistenza di un modello naturale. Quello che distingue i modelli naturali da quelli standard – come già sapeva Zermelo – è il fatto che l'operazione di passaggio all'insieme potenza è *assoluta* rispetto ad essi, nel senso che se $a \in V_\alpha$ (il livello α-esimo della gerarchia di von Neumann) allora l'insieme potenza di a considerato nell'universo V della metateoria coinciderà con l'insieme delle parti che sono elementi di V_α, nel caso ovviamente in cui V_α sia un modello di $\mathfrak{ZF}$.

In un senso quindi i modelli naturali sono quelli in cui il concetto di insieme potenza viene interpretato sul «vero» insieme potenza. Questo ad esempio non succede nel modello che già co-

nosciamo dei costruibili, in cui per ogni ordinale α, $L_\alpha \subseteq V_\alpha$ e l'inclusione in generale è propria, nel senso che non tutte le parti di un insieme risultano definibili. Ogni livello della gerarchia dei costruibili, in ogni caso, ci dà un'interpretazione standard e il confronto con la gerarchia di von Neumann ci permette di porre in luce una distinzione estremamente importante, che riguarda più in generale le classi che sono modelli di $\mathfrak{ZF}$. Gli ordinali che si ritrovano nella gerarchia dei costruibili e in quella di von Neumann sono gli stessi, e infatti, come si ricorderà, l'idea di Gödel era quella di considerare gli insiemi costruibili come gli insiemi predicativamente definibili una volta che si assumano come dato di partenza *tutti* gli ordinali, predicativi o no. Possiamo così dire che l'universo dei costruibili e quello di von Neumann hanno la stessa *altezza* ma non la stessa *larghezza* e questa distinzione si può facilmente estendere a strutture che non sono classi, prendendo come altezza il più piccolo ordinale che non compare nella nostra struttura standard transitiva.

Il problema che naturalmente sorgeva, una volta isolati i modelli naturali, era quello di determinare per quali α V_α è un modello di $\mathfrak{ZF}$. È facile vedere che per ogni $\alpha > 0$ risultano veri gli assiomi di estensionalità, di fondazione, di isolamento, dell'insieme vuoto, dell'unione e, banalmente, di scelta. Per soddisfare l'assioma della coppia e l'assioma della potenza occorre che α sia un ordinale limite, e questo è ovvio se si pensa che se $a \in V_\alpha$, esisterà un $\beta \leq \alpha$ tale che $a \subseteq V_\beta$; $\mathscr{P}(a)$ in generale apparterrà a $V_{\beta+1}$, così, se α non fosse limite (e quindi risultasse successore) si darebbero insiemi a la cui potenza non appartiene a V_α.

Infine è del tutto ovvio che, se vogliamo soddisfare l'assioma dell'infinito, occorre che α sia maggiore di ω. Il problema che rimane è allora quello dell'assioma di rimpiazzamento. È qui che interviene il concetto di cardinale inaccessibile: la possibilità che ci offre l'assioma di rimpiazzamento di fare riunioni su famiglie di insiemi indiciate da funzioni definite su insiemi, ci dice che α dovrà essere regolare nel senso sopra introdotto, cosicché possiamo concludere che un modello naturale V_α sarà modello di $\mathfrak{ZF}$ nel caso α sia un cardinale inaccessibile. Tarski introdusse il concetto di *universo* per indicare quegli insiemi transitivi che contengono l'ordinale ω, sono chiusi rispetto all'unione e l'insieme potenza e contengono l'immagine di tutti gli insiemi α ad essi appartenenti ri-

spetto a funzioni che sono loro sottoinsiemi, e dimostrò che gli universi coincidono con i V_κ dove κ è inaccessibile. Va detto però, come dimostrarono Montague e Vaught nel 1952, che gli universi non sono gli unici possibili modelli naturali di $\mathfrak{ZF}$, in quanto se V_α è un universo, si può provare che esiste un $\beta < \alpha$ tale che V_β è sottostruttura *elementare* di V_α.

Ci siamo limitati finora ai modelli di $\mathfrak{ZF}$, ma l'interesse dei ricercatori negli anni cinquanta non si restrinse a questa teoria ma coinvolse anche $\mathfrak{NBG}$ e $\mathfrak{MK}$. Che modelli avevano queste teorie? Se ci limitiamo a strutture standard transitive avremo che un'interpretazione per $\mathfrak{NBG}$ e $\mathfrak{MK}$ sarà data da una coppia $\langle M, C\rangle$ dove M sarà il dominio degli insiemi e C quello delle classi, con $C \subseteq \mathscr{P}(M)$. È ovvio che se $\langle M, C\rangle$ è modello di $\mathfrak{NBG}$, M determinerà un modello di $\mathfrak{ZF}$. Il problema è, al contrario, ottenere un modello di $\mathfrak{NBG}$ da un modello di $\mathfrak{ZF}$. Come Mostowski dimostrò nel 1950, se M è un modello di $\mathfrak{ZF}$, per ottenere un modello di $\mathfrak{NBG}$ basta considerare la famiglia C dei sottoinsiemi di M definibili, eventualmente con parametri, mediante formule del linguaggio. Avremo così un modello di $\mathfrak{NBG}$ ed è in questo modo che Mostowski provò, come ricordato, che $\mathfrak{NBG}$ è estensione conservativa di $\mathfrak{ZF}$. Il metodo non si applica nel caso di $\mathfrak{MK}$ che, come sappiamo, non è estensione conservativa di $\mathfrak{ZF}$, e non solo le strutture $\langle M, C\rangle$, dove M è modello di $\mathfrak{ZF}$ non danno modelli di $\mathfrak{MK}$ ma, come provò Shepherdson, questo non si dà mai, nel senso che *nessuna struttura* $\langle V_\alpha, C\rangle$, dove C è la famiglia dei definibili, può essere modello di $\mathfrak{MK}$. Questo ci conferma la natura impredicativa di $\mathfrak{MK}$, precisata da un ulteriore risultato di Shepherdson, che ci dice che se α è inaccessibile, modello di $\mathfrak{MK}$ sarà la struttura $\langle V_\alpha, V_{\alpha+1}\rangle$ dove appunto $V_{\alpha+1}$ contiene *tutte* le parti di V_α, anche quelle non definibili.

In tutti i discorsi sino ad ora abbiamo fatto riferimento da una parte ad un universo ambiente V che è quello della metateoria (e che costituisce un modello di $\mathfrak{ZF}$ più eventualmente altri assiomi, tipo l'esistenza di inaccessibili o l'assioma di scelta) dall'altra a insiemi in questo universo che danno origine a modelli standard o a modelli naturali. Trovare un insieme che sia modello significa in un certo senso trovare un insieme che «rifletta» proprietà dell'universo, chiaramente non tutte, ma quelle che bastano per essere modello di $\mathfrak{ZF}$. L'idea di considerare in generale questo tipo di

situazioni dette origine nei primi anni sessanta a ricerche interessanti di Bernays, Montague, Levy ed altri, tra cui Scott e Bruno Scarpellini, che collegavano l'esistenza di grandi cardinali alla possibilità di «riflettere» proprietà dell'universo in universi parziali. Per cominciare, precisiamo le idee ponendo per definizione che un insieme $X \in V$ *riflette* V *per* $\mathscr{A}$ quando

$$\forall x_1, \ldots, x_n \in X(\mathscr{A}(x_1, \ldots, x_n) \longleftrightarrow \mathscr{A}^X(x_1, \ldots, x_n)).$$

Ora il principio che Montague e Levy (simultaneamente a Scott e Scarpellini) dimostrarono nei primi anni sessanta è che sostanzialmente per ogni formula $\mathscr{A}$ esiste un universo parziale V_α che riflette l'universo totale relativamente ad essa nel senso che

$$\mathfrak{Z}\mathfrak{F} \vdash \forall \alpha \exists \beta > \alpha \forall x_1, \ldots, x_n \in V_\beta (\mathscr{A}(x_1, \ldots, x_n) \longleftrightarrow \mathscr{A}^{V_\alpha}(x_1, \ldots, x_n)).$$

La dimostrazione si basa sull'idea di eliminare i quantificatori esistenziali di una forma prenessa di $\mathscr{A}$ utilizzando funzioni di Skolem che ci portano a livelli più alti della gerarchia di von Neumann, sino a che, mediante un opportuno passaggio al limite, si raggiunga la stabilità e il livello V_α così ottenuto risulti chiuso rispetto alle funzioni di Skolem opportune. Per garantirsi questi livelli, è necessario utilizzare gli assiomi dell'infinito e di rimpiazzamento ed il fatto interessante è che *vale l'implicazione inversa*, cioè si può provare che il principio di riflessione sopra isolato è equivalente alla congiunzione dei due assiomi.

Questo fatto apriva una prospettiva che Bernays, Levy, e successivamente altri, non tardarono a sviluppare: l'idea cioè che mediante opportuni principi di riflessione, più forti di quello dato, sia possibile presentare sistematicamente assiomi dell'infinito sempre più forti che ci garantiscono l'esistenza di grandi cardinali. Vedremo nel prossimo capitolo come questa ricerca si sia sviluppata e come dopo i risultati di indipendenza di Cohen sia diventata urgente l'indagine e la classificazione sistematica dei grandi cardinali. Quello che ci preme sottolineare già adesso è la motivazione di fondo dei principi di riflessione, che si ricollega all'idea generatrice della gerarchia di von Neumann: l'universo degli insiemi non è una totalità compiuta, ma la possiamo approssimare sempre più da vicino postulando l'esistenza di universi parziali sufficiente-

mente grandi (indiciati da ordinali sufficientemente grandi) che ci permettono di riflettere parzialmente le proprietà dell'universo globale, che rimane inattingibile.

Lo studio dei principi di riflessione, come più in generale il confronto fra diversi modelli di $\mathfrak{ZF}$, portava in primo piano il problema di indagare la *assolutezza* delle diverse nozioni insiemistiche, vale a dire la loro invarianza rispetto al passaggio da un universo a un altro. Come si ricorderà, in un certo senso questo era stato, con Skolem, il punto di partenza dell'indagine sui modelli della teoria degli insiemi. Col suo teorema Skolem aveva provato la non assolutezza delle nozioni fondamentali riguardanti la teoria della cardinalità e aveva concluso che, essendo relative le nozioni insiemistiche, non aveva senso cercare di fondare la matematica su di esse. In questo modo, come Skolem aveva iniziato lo studio dei modelli della teoria degli insiemi, aveva nel contempo negato l'interesse della teoria stessa dal punto di vista fondazionale, ed era stato von Neumann il primo dopo di lui a riprendere questi temi in una prospettiva più positiva.

Di fatto fu Gödel che nei suoi lavori sul modello dei costruibili riaffrontò il problema dell'assolutezza delle nozioni insiemistiche, nella convinzione che la non assolutezza non fosse da esorcizzare, quanto da analizzare. Gödel considerava classi transitive e la variazione di significato che si verifica quando, supposto che la classe M sia sottoclasse della classe N, si interpreta una formula $\mathscr{A}$ in entrambi i modelli $<M, \in>$, $<N, \in>$. Dati $a_1, \ldots, a_n \in M$, può capitare che essi soddisfino la formula $\mathscr{A}$, in M ma non in N o viceversa. Nel caso la soddisfacibilità in M comporti quella in N per ogni $a_1, \ldots, a_n \in M$, diremo che la formula $\mathscr{A}$ è *persistente*, mentre diremo che è *assoluta* se vale anche il viceversa, cosicché dati $a_1, \ldots, a_n \in M$, essi godranno o meno della proprietà $\mathscr{A}$ a prescindere dall'universo M o N in cui li consideriamo «vivere». L'idea di assolutezza, così precisata, riflette la nozione intuitiva, in quanto sottolinea l'invarianza della proprietà rispetto al contesto in cui la consideriamo. Il fatto importante è che a questo punto possiamo classificare tutte le nozioni insiemistiche determinando se sono assolute o meno. Così ad esempio avremo che sono assolute l'intersezione, la riunione, l'insieme coppia, ma non l'insieme potenza; saranno assoluti i singoli ordinali ma non i cardinali, ecc. Fatto importante, come Gödel dimostrò, la proprietà $L(x)$, «essere

costruibile», risulta essere assoluta, come pure la relazione di verità $\models$ che abbiamo a suo tempo definito.

Dalla teoria generale dei modelli sappiamo che quando consideriamo due strutture $\mathfrak{A}$ e $\mathfrak{B}$ di cui $\mathfrak{A}$ sottostruttura di $\mathfrak{B}$, il fatto che una formula sia preservata o assoluta è legata alla sua forma sintattica. Così sono assolute le formule prive di quantificatori e preservate le formule Σ_1. È questa considerazione che portò Levy, nel 1965, a definire esplicitamente una gerarchia di formule, nel linguaggio della teoria degli insiemi, che si basava sulla complessità dei prefissi, in stretta analogia ad esempio con la gerarchia aritmetica.

Levy considerava le formule a meno di equivalenza logica rispetto a sottoteorie di $\mathfrak{ZF}$ e la novità più importante è che, nel caso degli insiemi, assolute non sono solo le formule prive di quantificatori, ma anche quelle in cui i quantificatori sono limitati, della forma cioè $\exists x \in y$, $\forall x \in y$. Questo dipende dal fatto che le classi che consideriamo sono transitive; se allora definiamo Δ_0^0-formule quelle che non contengono quantificatori che non siano limitati, potremo prolungare la gerarchia nel modo solito ottenendo Σ_n^1-e Π_n^1-formule, a seconda che si aggiungano quantificatori esistenziali o universali. La situazione generale dal punto di vista della teoria dei modelli è stata esaminata da Feferman nel 1968, considerando la relazione di *estensione finale* tra modelli, ma quello che ci interessa rilevare in questo contesto è che la gerarchia di Levy ha aperto la strada ad uno studio sistematico della complessità delle formule insiemistiche che in un certo modo ci permette di graduare la non assolutezza delle nozioni che consideriamo.

Le considerazioni di assolutezza ci permettono di ottenere informazioni importanti sui rapporti tra costruibili e gerarchia di von Neumann. Come Kreisel sottolineò nel 1956, il fatto che ω sia assoluto comporta che se chiamiamo *formule aritmetiche* quelle i cui quantificatori sono ristretti a ω, avremo che una formula aritmetica è vera in V se e solo se è vera nell'universo dei costruibili e questo ci permette di concludere che se dimostriamo una formula aritmetica $\mathscr{A}$ utilizzando l'assioma di costruibilità $V = L$ oltre agli assiomi di $\mathfrak{ZF}$, potremo dimostrare la stessa formula dal solo $\mathfrak{ZF}$. Possiamo così affermare che quando utilizziamo *AS* o *IGC* (che come sappiamo sono conseguenze di $V = L$) per dimostrare una formula aritmetica, l'uso di questi assiomi non è essenziale ed è

eliminabile, nel senso che la formula stessa sarà dimostrabile dal solo $\mathfrak{ZF}$. Dal punto di vista metamatematico la cosa è estremamente interessante in quanto, ad esempio, ci permette di utilizzare, poniamo, modelli saturi per provare fatti sintattici (e quindi, via gödelizzazione, aritmetici) utilizzando eventualmente *IGC*. Questo si verifica ad esempio per alcune dimostrazioni del teorema di interpolazione ed altri risultati analoghi; l'importante è che, per quanto detto, si ha la sicurezza che l'ipotesi in linea di principio è eliminabile.

Queste considerazioni hanno un altro risvolto, che riguarda la stessa teoria degli insiemi: quali enunciati provabili assumendo l'ipotesi di costruibilità sono dimostrabili anche *senza* di essa? Nel 1959 Shoenfield riuscì a provare che all'interno della gerarchia aritmetica solo a partire dagli enunciati Σ^1_3 in poi si potevano trovare formule che godessero di questa proprietà, in quanto il risultato di Kreisel sopra citato si estendeva sia alle Σ^1_2 sia alle Π^1_2.

Ma le indagini di questo tipo non tardarono a mettere in evidenza un fatto più decisivo, che riguardava questa volta non solo i costruibili, ma l'uso stesso dei modelli interni per le dimostrazioni di indipendenza. Furono le ricerche di J. Shepherdson dei primi anni cinquanta sui modelli interni che misero in luce i limiti che questi presentavano nello studio delle questioni di indipendenza. I risultati di Shepherdson riguardavano il concetto di *modello minimale*. Assumiamo che $\mathfrak{ZF}$ abbia un modello transitivo. Possiamo dimostrare, ricorrendo all'assolutezza della nozione di costruibilità, che esisteranno ordinali α per cui L_α è anch'esso modello di $\mathfrak{ZF}$ e si può pure provare che il più piccolo di tali ordinali sarà numerabile: indichiamolo con α_0. Avremo allora che L_{α_0} è un modello minimale di $\mathfrak{ZF}$ nel senso che sarà incluso in ogni modello transitivo di $\mathfrak{ZF}$. La cosa è ovvia in quanto, ancora, per l'assolutezza della proprietà di costruibilità, se β è l'ordinale massimo che occorre in una classe transitiva X modello di $\mathfrak{ZF}$, avremo che L_β coincide con la classe dei costruibili in X e sarà modello di $\mathfrak{ZF}$; d'altra parte, per la minimalità di α_0, avremo che $\alpha_0 \leq \beta$ e quindi $L_{\alpha_0} \subseteq X$.

Questo significa, dal punto di vista delle dimostrazioni di indipendenza, che non è possibile utilizzare il metodo dei modelli interni per ottenere contromodelli di *AS*, *IGC*, o $V = L$. I modelli interni, come sappiamo, sono determinati da classi definibili da for-

mule $\mathscr{A}(x)$. Per falsificare *AS* o *IGC*, una classe *D* definita da $\mathscr{A}(x)$ deve falsificare $V = L$ che implica entrambi gli assiomi. La formula $\mathscr{A}(x)$ induce anche in L_α un modello interno $\mathfrak{M} \subseteq L_\alpha$. Per minimalità, $\mathfrak{M} = L_\alpha$ e così non può essere che $\mathfrak{M}$ falsifichi l'assioma di costruibilità, per l'assolutezza di *L*.

Contrariamente a quanto ci si potrebbe aspettare, il risultato di Shepherdson sul modello minimale non fu visto subito in quest'ottica e solo quando Cohen lo riscoprì nei primi anni sessanta fu chiaro che per provare l'indipendenza degli assiomi sopraddetti occorrevano metodi di costruzione completamente nuovi, in particolare tecniche che permettessero di costruire *estensioni* piuttosto che *sottostrutture*.

Prima di passare a parlare dei metodi che Cohen introdusse allo scopo, conviene però soffermarsi ancora un attimo sugli insiemi costruibili e su un altro contributo importante negli anni precedenti il lavoro di Cohen: la definizione di «costruibilità relativa» data da A. Hajnal nel 1956 e da Levy nel 1960.

L'idea di fondo del concetto di costruibilità relativa è semplicemente quella di definire, associata ad ogni insieme transitivo *X*, una gerarchia $L_\alpha(X)$ ponendo $L_0(X) = X$ e procedendo poi per gli ordinali successori e limiti precisamente come nell'usuale costruibilità. Gli insiemi della classe $L(X)$ che si ottiene facendo la riunione su tutti gli ordinali si chiamano insiemi *costruibili da X* e intuitivamente si possono vedere come gli insiemi predicativamente definibili ottenuti assumendo come dati non solo gli ordinali ma lo stesso insieme *X*. Anche in questo caso la relazione $y = L_\alpha(X)$ sarà assoluta, e sfruttando l'ipotesi che *X* sia transitivo avremo anche che ogni $L_\alpha(X)$ è transitivo. Si può così provare, in stretta analogia col caso assoluto, che otterremo in questo modo modelli transitivi di $\mathfrak{Z}\mathfrak{F}$, con l'aggiunta significativa che in certi casi alcune delle proprietà di *X* vengono ereditate dalla gerarchia: ad esempio si può provare che se *X* ha un buon ordinamento in $L(X)$, quest'ultimo soddisferà l'assioma di scelta.

La situazione cambia nel caso di *IGC*, perché ora *X* ha già una cardinalità, su cui non possiamo agire, sicché senza ulteriori assunzioni si possono provare risultati di carattere più limitato, ad esempio che nell'ipotesi $V = L(X)$, la *IGC* vale per tutti i cardinali κ maggiori o uguali alla cardinalità di *X*. Le applicazioni del concetto di costruibilità relativa che diedero Levy e Hajnal sono nu-

merose, ma forse più interessante è la prospettiva che tale concetto apriva sulla possibilità di costruire estensioni di modelli dati a partire da insiemi aggiuntivi, «chiudendo» rispetto alla definibilità. Sarà questa la via che seguirà Cohen.

L'importanza dei lavori di Cohen infatti non sta solo nell'aver risolto un problema vecchio di quasi un secolo, ma nella tecnica completamente nuova che presentava per costruire modelli di $\mathfrak{ZF}$, tecnica le cui applicazioni non tardarono a rivelarsi estremamente feconde in contesti svariatissimi. L'originalità di questi metodi stupì gli stessi logici, che in buona parte – con l'eccezione sicura di Gödel – erano giunti alla conclusione che il problema dell'indipendenza di AS e IGC trascendeva i mezzi disponibili. Non è un caso che Cohen non fosse un logico di formazione, bensì un analista, che dovette scontrarsi con lo scetticismo di alcuni fra i logici più in vista.

Come detto sopra parlando del modello minimale, lo scoglio centrale da superare, una volta che si voglia dimostrare l'indipendenza di AS e IGC, è costituito dal fatto di non poter ricorrere ai modelli interni, e di dover quindi costruire *estensioni* di modelli. Possiamo illustrare l'idea di Cohen limitandoci al caso dell'assioma di costruibilità, in quanto, come sappiamo, ogni contromodello di AS e IGC dovrà essere anche un contromodello di $V = L$. Per falsificare tale assioma è sufficiente costruire un modello in cui si trova un reale (vale a dire un sottoinsieme di N) che non è costruibile. L'idea naturale sarebbe quindi quella di partire da un modello $\mathfrak{M}$ standard transitivo che – come Cohen dimostra – possiamo assumere addirittura numerabile, e «aggiungergli» un reale non costruibile a, chiudendo poi l'universo in modo da avere ancora un modello di $\mathfrak{ZF}$.

L'ostacolo a questa procedura è dato dal fatto che in generale non siamo in grado di controllare il comportamento di queste estensioni, ed in particolare non è detto a priori che se a non è costruibile nel modello di partenza non lo diventi nel modello esteso. Il passo necessario da compiere per avere questo è quello di garantirsi che il nuovo modello contenga gli stessi ordinali del modello di partenza. Per l'assolutezza della relazione «$y = L_\alpha$», sappiamo infatti che se due modelli standard hanno gli stessi ordinali, avranno gli stessi costruibili. È a questo punto che interviene in modo decisivo il concetto di *forcing* mediante il quale Co-

hen è in grado di controllare le proprietà delle estensioni che costruisce.

Il *forcing* è una relazione sintattica che costituisce una sorta di definizione di verità relativa a pezzi *finiti* di informazione riguardanti gli elementi dell'insieme a, il nuovo reale non costruibile, che vogliamo aggiungere. È chiaro infatti che le proprietà del modello $\mathfrak{N}$ che costruiamo dipenderanno in qualche modo dalle proprietà dell'insieme a che aggiungiamo. In generale infatti, per quanto detto sopra, chiudere l'universo in modo da ottenere un modello significa considerare i costruibili in a, ottenibili utilizzando solo il segmento degli ordinali presenti nel modello di partenza. La strategia dovrà essere allora quella di impedire che a contenga informazioni speciali su $\mathfrak{M}$, il modello da cui siamo partiti, perché, come si può provare, se si prendono a «troppo» informativi c'è il rischio che il nostro modello $\mathfrak{N}$ contenga elementi indesiderati. Ad esempio, se α_0 è il più piccolo ordinale numerabile non occorrente in M, associato ad esso sarà un sottoinsieme di N^2 che – codificando le coppie – possiamo identificare direttamente con un sottoinsieme di N. Ora se prendiamo come a questo sottoinsieme, è facile verificare che ogni modello di $\mathfrak{ZF}$ contenente a dovrà contenere α_0, cosicché il nostro modello $\mathfrak{N}$ dovrebbe avere α_0 fra i suoi ordinali per essere un modello: contro il fatto che noi non vogliamo nuovi ordinali nel nuovo modello.

Si tratta allora di costruire $\mathfrak{N}$ imponendo che in esso a si comporti in modo «generico» rispetto a $\mathfrak{M}$, non abbia in altre parole proprietà particolari che lo distinguano «troppo» da altri elementi in $\mathfrak{M}$. È chiaro d'altra parte che ci sono affermazioni su a che in nessun modo possiamo interpretare come proprietà generali in senso assoluto, per esempio enunciati del tipo $n \in a$ o $n \notin a$ dove n è un naturale. In certo modo quindi la genericità deve essere compatibile con questo tipo di informazioni e l'idea fondamentale di Cohen è appunto quella di considerare gli insiemi finiti coerenti di enunciati di questo tipo come *condizioni* su a e più in generale sul modello che vogliamo costruire. Quest'ultimo si otterrà considerando le formule «forzate» non da una singola condizione, ma da un insieme di condizioni sufficienti per determinare univocamente a. L'importante è che questi insiemi di condizioni, se determinano completamente a, impongono la verità di date formule nel modello $\mathfrak{N}$ in modo «generico», sfruttando *solo* singole condizioni. La

«genericità» si otterrà se chiamiamo «insiemi generici di condizioni» quelle famiglie in grado di determinare univocamente un sottoinsieme di N e definiamo la relazione di *forcing* in modo che una condizione P «forza» un enunciato $\mathscr{A}$ se e solo se questo avviene rispetto ad *ogni* generico che contenga P.

La descrizione ora data è abbastanza vaga, ma sottolinea quelli che ci sembrano i punti essenziali del metodo proposto da Cohen e l'idea intuitiva su cui esso si basa. Sul piano più tecnico vanno fatte alcune precisazioni. La prima è che Cohen utilizza come modello standard transitivo di partenza quel modello minimale $\mathfrak{M}$, numerabile, la cui esistenza aveva precedentemente dimostrato. Il modello $\mathfrak{N}$ si otterrà da $\mathfrak{M}$ considerando i costruibili in un dato insieme a non appartenente a $\mathfrak{M}$. Quali a si debbano scegliere dipende dalla nostra definizione di *forcing* e quindi dal linguaggio in cui ci muoviamo. Quello che Cohen introduce è un linguaggio ramificato in cui sostanzialmente si hanno quantificatori limitati del tipo $\forall_\alpha x$ o $\exists_\alpha x$ per ogni $\alpha < \alpha_0$, il più piccolo ordinale non appartenente al modello minimale. L'idea di utilizzare un linguaggio ramificato con ordinali $\beta < \alpha_0$ è dovuta al fatto che per definizione il modello $\mathfrak{N}$ che vogliamo ottenere è dato dagli insiemi costruibili in a e quindi da insiemi della forma

$$\{x \mid \mathscr{A}(x)\}$$

dove $\mathscr{A}$ è un enunciato limitato, in cui i quantificatori vengono naturalmente interpretati come relativizzati ai livelli della gerarchia dei costruibili. Per Cohen il modello $\mathfrak{N}$ si dovrà ottenere sostanzialmente come i modelli canonici costruiti da insiemi di formule T completi e ricchi nella dimostrazione di completezza di Henkin, vale a dire quozientando l'insieme dei termini rispetto alla equivalenza $\approx$ definita ponendo $t \approx t'$ se e solo se $(t = t') \in T$, e definendo l'interpretazione del simbolo $\in$ in modo tale che

$$[t] \in [t'] \text{ se e solo se } (t \in t') \in T$$

dove $[t]$ è la classe di equivalenza di t.

Il ruolo del *forcing* è appunto quello di permetterci di costruire un insieme T che descrive un modello $\mathfrak{N}$ la cui costruzione sarà come quella di sopra.

Il *forcing* viene definito tra *condizioni*, insiemi finiti di enunciati del tipo $n \in \bar{a}$, $n \notin \bar{a}$, dove $\bar{a}$ è una costante individuale che sarà il nome del nostro insieme a, e formule. La collezione C di tutte le condizioni costituisce un insieme ordinato $(C, \subseteq)$ rispetto all'inclusione insiemistica e la relazione di *forcing* è definita facendo riferimento a questa struttura d'ordine. Intuitivamente l'idea è che quando si tratta di stabilire se P forza l'enunciato $\neg \mathscr{A}$, non ci si deve limitare solo alle informazioni contenute in P ma occorre tener conto di tutte le possibili condizioni Q che la estendono: è questo che ci garantisce che se P forza $\mathscr{A}$ e Q estende P, allora anche Q forzerà $\mathscr{A}$ per ogni formula $\mathscr{A}$.

Originariamente Cohen definiva la relazione di *forcing* solo per formule prenesse; in seguito la definizione fu modificata da Scott, che mostrò il ruolo centrale della clausola sulla negazione. Se – senza entrare in ulteriori dettagli – indichiamo con S_β l'insieme dei nomi per elementi costruiti al livello β-esimo della gerarchia (e che immaginiamo associati biunivocamente a formule $\mathscr{A}(x)$ che contengono nomi e quantificatori di livello più basso di β), possiamo allora definire come segue la relazione «P forza $\mathscr{A}$» ($P \Vdash \mathscr{A}$) per formule $\mathscr{A}$ limitate, lasciando per ultimo il caso delle formule atomiche:

1) se $\mathscr{A} = (\mathscr{A}_1 \wedge \mathscr{A}_2)$ $[\mathscr{A}_1 \vee \mathscr{A}_2]$, allora $P \Vdash \mathscr{A}$ se e solo se $P \Vdash \mathscr{A}_1$ e [o] $P \models \mathscr{A}_2$

2) se $\mathscr{A} = (\mathscr{A}_1 \rightarrow \mathscr{A}_2)$, allora $P \Vdash \mathscr{A}$ se e solo se $P \Vdash \neg \mathscr{A}_1$ oppure $P \Vdash \mathscr{A}_2$

3) se $\mathscr{A} = (\neg \mathscr{A}_1)$, allora $P \Vdash \mathscr{A}$ se e solo se per nessuna estensione Q di P, $Q \Vdash \mathscr{A}_1$

4) se $\mathscr{A} = (\exists_\beta x \mathscr{A}_1(x))$, allora $P \Vdash \mathscr{A}$ se e solo se esiste una $c \in S_\alpha$ per $\alpha < \beta$ tale che $P \Vdash \mathscr{A}_1(\bar{c})$

5) se $\mathscr{A} = (\forall_\beta x \mathscr{A}_1)$, allora $P \Vdash \mathscr{A}$ se e solo se $P \Vdash \neg \exists_\beta x \neg \mathscr{A}_1(x)$.

Rimane il caso delle formule atomiche che possono essere di diverso tipo. Se la formula è del tipo $(n \in \bar{a})$ (dove, come si ricorde-

rà, $\bar{a}$ è il nome del reale non costruibile che vogliamo aggiungere a $\mathfrak{M}$) allora porremo che $P \Vdash \mathscr{A}$ se e solo se $(n \in \bar{a}) \in P$. Le altre formule atomiche sono del tipo $\bar{c} \in \bar{c}'$ e $\bar{c} = \bar{c}'$ dove $c \in S_\alpha$ e $c' \in S_\beta$. Senza entrare in dettagli, il principio da seguirsi in questi due casi è il seguente: i termini $\bar{c}$ e $\bar{c}'$ denoteranno elementi del modello $\mathfrak{N}$ e come sappiamo, se vogliamo che risulti vero l'assioma di estensionalità, occorre che gli insiemi denotati dai termini siano uguali (quindi che sia forzata da P la formula $(\bar{c} = \bar{c}')$) nel caso tutti gli elementi di c siano elementi di $\bar{c}'$ e viceversa. Gli elementi di $\bar{c}$ e $\bar{c}'$ apparterranno a livelli inferiori a quelli di c e c', quindi inferiori a $\gamma = \max(\alpha, \beta)$. Di conseguenza la clausola $P \Vdash \bar{c} = \bar{c}'$ si ricondurrà al caso

$$P \Vdash \forall_\gamma\, x(x \in \bar{c} \longleftrightarrow x \in \bar{c}')$$

In modo analogo si tratta il caso $\bar{c} \in \bar{c}'$ *e si impone che ogni* $\bar{c}$ sia interpretato come l'insieme degli x che soddisfano la formula associata $\mathscr{A}(x)$.

Importante è osservare che la correttezza della definizione della relazione di *forcing* dipende da un'opportuna assegnazione di ranghi ordinali alle formule ed è quindi una definizione induttiva la quale – è questo un aspetto essenziale di tutta la costruzione – non solo risulta assoluta, ma è *esprimibile* all'interno di $\mathfrak{ZF}$ (più precisamente, lo è *fissata* la condizione P). Con naturali modificazioni si può definire la relazione di *forcing* per enunciati illimitati (i cui quantificatori non siano cioè limitati).

Rimane ora il passo finale: la costruzione della struttura $\mathfrak{N}$, la dimostrazione che essa è modello di $\mathfrak{ZF}$ e la verifica che, in $\mathfrak{N}$, $V = L$ risulta falsificata dal nuovo reale a. Il concetto fondamentale è a questo punto quello di *successione completa* di condizioni. Si dice che una successione $\{P_n\}$ di condizioni è completa, se per ogni n e ogni formula $\mathscr{A}$, limitata o illimitata che sia, esiste un $k \geq n$ tale che

$$P_k \Vdash \mathscr{A} \quad \text{oppure} \quad P_k \Vdash \neg \mathscr{A}.$$

È immediato che per definizione se $\{P_n\}$ è una successione completa, l'insieme T delle formule forzate da almeno una condizione della successione conterrà per ogni enunciato $\mathscr{A}$ o $\neg \mathscr{A}$. Ma si può provare di più. Per definizione, per ogni P e per ogni $\mathscr{A}$ non si

può avere simultaneamente che $P \Vdash \mathscr{A}$ e $P \Vdash \neg \mathscr{A}$ e, ancora, si può provare che

$$\text{se } P \Vdash \mathscr{A} \text{ e } P \subseteq Q \text{ allora } Q \Vdash \mathscr{A}.$$

Da qui scende che T sarà un insieme completo nel senso che per ogni enunciato $\mathscr{A}$ conterrà $\mathscr{A}$ o $\neg \mathscr{A}$ ma non entrambi e questo ci mostra allora in che senso la costruzione «alla Henkin» di $\mathfrak{N}$ sia possibile utilizzando il *forcing*.

Si tratterà ora di dimostrare che *esiste* una successione completa, e questo richiede due fatti. Il primo è una ovvia proprietà generale del *forcing* che ci dice che per tutte le P e tutte le $\mathscr{A}$

$$\text{esiste una } Q \supseteq P \text{ tale che } Q \Vdash \mathscr{A} \text{ oppure } Q \Vdash \neg \mathscr{A}.$$

Il secondo chiama in causa – ed è la sola volta che ciò avviene – la *numerabilità* del nostro modello minimale $\mathfrak{M}$. Se $\mathfrak{M}$ è numerabile, tutti gli enunciati del linguaggio lo saranno, così che con una semplice procedura induttiva possiamo definire la nostra successione completa sfruttando la proprietà precedente.

C'è un'osservazione da fare. Quando diciamo che esiste una successione completa, questo non significa che tale successione appartenga a $\mathfrak{M}$, anche se ogni singola condizione è in $\mathfrak{M}$; semplicemente, la successione apparterrà all'universo nel quale ci muoviamo ed è essenziale che così sia, in quanto se la successione appartenesse ad $\mathfrak{M}$, la nostra costruzione potrebbe risultare non adeguata.

Passando al modello canonicamente associato all'insieme T individuato dalla successione completa che abbiamo visto esistere, otteniamo una struttura che rende vere *esattamente* le formule forzate da almeno una condizione della successione. Sfruttando ora la definibilità della relazione di *forcing* all'interno di $\mathfrak{ZF}$ per ogni condizione P, possiamo provare che $\mathfrak{N}$ sarà un modello di $\mathfrak{ZF}$ ed è essenziale in questo fatto che la verità in $\mathfrak{N}$ sia stata definita ricorrendo a quantità *finite* di informazione, cioè a singole condizioni. È anche chiaro in che senso a e $\mathfrak{M}$ sono *generici*; supponiamo infatti che $\mathfrak{M} \Vdash \neg\neg \mathscr{A}$; esisterà allora un P per cui $P \Vdash \neg\neg \mathscr{A}$. Come sarebbe emerso più avanti dire che $P \Vdash \neg\neg \mathscr{A}$ è lo stesso che dire che $\mathscr{A}$ è vera in *tutti* i modelli associati (nel senso di sopra) a tut-

te le successioni complete che contengono P. La relazione di *forcing* non discrimina tra successioni complete e il fatto che $P \Vdash \neg\neg \mathscr{A}$ non comporta la scelta di una particolare successione: *localmente* la verità in $\mathfrak{N}$ dipende da singole condizioni, mentre è solo *globalmente* che dipende da una specifica successione completa. È sfruttando questa genericità che si verifica che a non è costruibile e si può concludere che $\mathfrak{N}$ falsifica l'assioma di costruibilità.

Se la costruzione di $\mathfrak{N}$ ci ha permesso di vedere nelle sue linee essenziali il funzionamento del metodo del *forcing*, per ottenere modelli che falsifichino AS e IGC occorre articolare maggiormente la costruzione di partenza. Come si può verificare infatti $\mathfrak{N}$ pur falsificando $V=L$, soddisfa tanto AS che IGC. Il metodo che Cohen segue per falsificare AS poggia, oltre che sul *forcing*, su delle tecniche che hanno una lunga storia e che nel loro nucleo centrale risalgono agli anni venti, quando Fraenkel, nel lavoro del 1922 a suo tempo citato, fornì un metodo per provare l'indipendenza di AS dagli altri assiomi di Zermelo.

Come del resto anche il sistema di Zermelo, il sistema modificato di Fraenkel differiva però dall'attuale teoria $\mathfrak{ZF}$ per l'esplicita ammissione nell'universo di *Urelemente* e, dal nostro punto di vista, si può assimilare al sistema $\mathfrak{ZFA}$ ($\mathfrak{ZF}$ con atomi) introdotto in tempi più vicini a noi; $\mathfrak{ZFA}$ nel linguaggio contiene una costante predicativa unaria A («essere atomo») e identifica gli insiemi come non-atomi, riformulando opportunamente gli assiomi di estensionalità e fondazione.

L'idea di fondo di Fraenkel era di considerare modelli di $\mathfrak{ZFA}$ in cui esiste un insieme A di atomi numerabile e di provare che questo stesso insieme A non ammette un buon ordinamento. Strumento centrale di questa analisi era il concetto di *automorfismo* (isomorfismo di una struttura con se stessa) dell'universo insiemistico, concetto che sta alla base della riformulazione del metodo nel contesto della teoria dei gruppi, presentata da Mostowski nel 1938 e successivamente perfezionata da Specker nel 1957. La tecnica è estremamente diretta e costituisce il prototipo del metodo che poi seguirà Cohen.

Per cominciare possiamo osservare che, in forza dell'assioma di fondazione opportunamente modificato per la presenza degli atomi, è possibile, anche per i modelli di $\mathfrak{ZFA}$, costruire una gerarchia analoga a quella di von Neumann, il cui primo livello $V_0(A)$ coin-

cide appunto con l'insieme A degli atomi. Quello che si fa con l'insieme A degli atomi si può fare a partire da qualunque insieme S dell'universo e se S coincide con l'insieme vuoto avremo una gerarchia modello di $\mathfrak{ZF}$ che chiameremo il *nucleo* del nostro universo $V(A)$. Tutti gli ordinali, per definizione, apparterranno al nucleo, e il buon ordinamento di A, la cui esistenza vogliamo negare, sarà un'applicazione che ci porta da un ordinale nel nucleo sull'insieme A. L'idea di fondo del metodo di Fraenkel-Mostowski è appunto che gli assiomi di $\mathfrak{ZFA}$ non ci permettono di distinguere un atomo da un altro, cosicché A non sarà bene ordinabile.

È per rendere l'idea di indistinguibilità che si ricorre al concetto di automorfismo: ogni permutazione π dell'insieme A induce un automorfismo dell'intero universo $V(A)$ nel senso che si prolungherà ad una permutazione $\bar{\pi}$ di $V(A)$ che conserva $\in$ nel senso che

$$x \in y \text{ se e solo se } \bar{\pi}(x) \in \bar{\pi}(y).$$

Consideriamo ora un gruppo $\mathfrak{G}$ di permutazioni di A e sia F un filtro di sottogruppi di $\mathfrak{G}$ chiuso rispetto al passaggio ai gruppi coniugati e tale che

$$\text{per ogni } a \in A, \{\pi \in \mathfrak{G} \mid \bar{\pi}(a) = a\} \in F.$$

Chiamiamo *normale* un filtro di questo tipo ed in generale, per ogni elemento x dell'universo, poniamo

$$\mathrm{sim}_{\mathfrak{G}}(x) = \{\pi \in \mathfrak{G} \mid \bar{\pi}(x) = x\}.$$

Fissati $\mathfrak{G}$ e F, diremo *simmetrico* un elemento x dell'universo tale che il suo gruppo di simmetria $\mathrm{sim}_{\mathfrak{G}}(x) \in F$.

Chiameremo poi *ereditariamente simmetrici* quegli elementi che, oltre ad essere simmetrici, hanno anche elementi ereditariamente simmetrici. La loro collezione S verrà detta un *modello di permutazione*, ed è facile verificare che costituirà un modello transitivo di $\mathfrak{ZFA}$. L'uso dei modelli di permutazione per dimostrare l'indipendenza di principi legati all'assioma di scelta si basa essenzialmente sull'individuazione di opportuni gruppi e filtri usando opportune famiglie di sottoinsiemi dell'insieme A degli atomi. Fissato un gruppo $\mathfrak{G}$ di permutazioni di A, diremo che una famiglia $\mathscr{I}$ di parti

di A è un *ideale normale,* se è un ideale che contiene per ogni $a \in A$ il singleton di a ed inoltre

$$\text{se } \pi \in \mathfrak{G} \text{ e } E \in \mathcal{I}, \text{ allora } \pi(E) \in \mathcal{I}.$$

Dato ora un elemento x sia

$$\text{fix}_{\mathfrak{G}}(x) = \{\pi \in \mathfrak{G} \mid \overline{\pi}(y) = y\} \text{ per ogni } y \in x.$$

Ognuno di questi insiemi è un sottogruppo di $\mathfrak{G}$ e la loro totalità, al variare dei sottoinsiemi appartenenti a $\mathcal{I}$, genera un filtro normale che dà origine ad un modello di permutazione $\mathfrak{S}$ in cui un elemento x è *simmetrico* se e solo se esiste un $E \in \mathcal{I}$ tale che

$$\text{fix } \mathfrak{G}(E) \subseteq \text{sim}(x).$$

Un'ultima osservazione: se $\mathfrak{S}$ è un modello di permutazione, conterrà tutti gli elementi del nucleo e se – come faremo – partiamo da modelli di $\mathfrak{ZFA} + AS$, AS varrà nel nucleo così che ogni suo elemento sarà bene ordinato. Ne scende che preso $x \in \mathfrak{S}$, esso potrà essere bene ordinato se e solo se è in corrispondenza biunivoca con un elemento del nucleo. A questo punto è immediato verificare che

$$\mathfrak{S} \models (x \text{ è bene ordinabile}) \quad \text{se e solo se} \quad \text{fix}(x) \in F.$$

A questo punto abbiamo la chiave per costruire i modelli opportuni. Il modello fondamentale costruito da Fraenkel partiva da universi con insiemi di atomi A numerabili e considerava come $\mathcal{I}$ l'insieme delle parti *finite di A*. Risultava immediato che fix(A) *non* apparteneva al filtro individuato da $\mathcal{I}$ cosicché non era bene ordinabile, falsificando l'assioma di scelta. Con la stessa tecnica Fraenkel provava che in $\mathfrak{ZFA}$ non era dimostrabile AS per famiglie numerabili di coppie e Mostowski provava che AS in $\mathfrak{ZFA}$ era indipendente dal principio dell'ordinamento, per cui ogni insieme ammette un ordinamento lineare.

Le stesse tecniche furono utilizzate ampiamente negli anni cinquanta da altri ricercatori, e sempre in questo periodo Mendel-

son, Shoenfield, Specker e altri cercarono di estendere il metodo a universi *senza* atomi indebolendo l'assioma di fondazione. La speranza era di giungere così a provare l'indipendenza di AS da $\mathfrak{ZF}$. L'ostacolo principale in questa estensione stava nel fatto che se non si ammettono atomi e ci si muove in modelli standard transitivi, non ci sono automorfismi propri. Il problema è di trovare un modo di indurne: fu questo che riuscì a Cohen combinando la tecnica del *forcing* con la costruzione di insiemi generici.

L'idea di Cohen è di prendere insiemi generici al posto degli atomi e di applicare a questi il metodo delle permutazioni. Si tratta allora di costruire un'estensione $\mathfrak{N}_1$, del modello minimale aggiungendo (come nei modelli di Fraenkel) un'infinità numerabile di nuovi «atomi», cioè di insiemi generici $\{a_i\}$ con $i \in$ N. Cohen introduceva un opportuno linguaggio in cui disponeva di nomi $\bar{a}_i$ per questi insiemi che dovevano essere – come nel modello utilizzato per falsificare l'assioma di costruibilità – sottoinsiemi di N e considerava come condizioni gli insiemi coerenti finiti di formule del tipo $n \in \bar{a}_i$ e $n \notin \bar{a}_i$.

L'ordine fra le condizioni era dato dall'inclusione insiemistica e la definizione di *forcing* estendeva in modo ovvio quella data precedentemente modificando le clausole sulle atomiche. Dovendo procedere sul piano sintattico nella costruzione del modello desiderato, Cohen, in stretta analogia con quanto fatto da Fraenkel per il suo modello fondamentale, considerava il gruppo $\mathfrak{G}$ delle permutazioni π di N che differiscono dall'identità per un numero *finito* di argomenti e faceva agire $\mathfrak{G}$ sugli insiemi S_α di nomi del suo linguaggio imponendo per ogni $\pi \in \mathfrak{G}$:

$$\pi(\bar{a}_i) = \bar{a}_{\pi(i)} \text{ e } \pi(A) = A$$

dove A è il nome introdotto per indicare quella che sarà la collezione degli insiemi generici a_i ($i \in$ N). A questo punto è chiaro come anche ad ogni formula $\mathscr{B}$ del linguaggio si possa associare, mediante una permutazione $\pi \in \mathfrak{G}$, una formula $\pi(\mathscr{B})$ che si ottiene sostituendo ai nomi che occorrono in $\mathscr{B}$ i nomi che π associa ad essi. Il fatto importante è che si può provare che per ogni condizione P ed ogni formula $\mathscr{A}$

$$P \Vdash \mathscr{A} \text{ se e solo se } \pi(P) \Vdash \pi(\mathscr{A})$$

così che – in certo senso – ogni $\pi \in \mathfrak{G}$ conserva la relazione di *forcing* ed è in virtù di questo che, una volta scelta una successione completa e ottenuto il modello associato $\mathfrak{N}_1$, possiamo considerare quest'ultimo come una sorta di modello di permutazione in cui i generici fanno da atomi e le $\pi \in \mathfrak{G}$ sono automorfismi rispetto ai quali gli elementi di $\mathfrak{N}$ risultano ereditariamente simmetrici. Cohen poteva così provare che R (l'insieme dei reali) in $\mathfrak{N}_1$ non risultava bene ordinato (si ricordi che gli insiemi generici aggiunti sono «reali») e che quindi l'assioma di scelta risultava falsificato.

Questo risultato veniva presentato da Cohen in un «preprint» dal titolo *The Independence of the Axiom of Choice* (*L'indipendenza dell'assioma di scelta*) nel 1963 e solo successivamente, insieme ad altri risultati, venne esposto con tutti i dettagli nel libro *Set Theory and the Continuum Hypothesis* (*Teoria degli insiemi e ipotesi del continuo*) pubblicato nel 1965. Già nel preprint, accanto alla dimostrazione della indipendenza di $V = L$ da $\mathfrak{Z}\mathfrak{F} + AS$ e IGC e di quella di AS e della sua versione numerabile per coppie di funzioni reali, Cohen presentava il risultato che costituisce l'applicazione più clamorosa del suo metodo: la dimostrazione dell'indipendenza dell'ipotesi del continuo (IC) da $\mathfrak{Z}\mathfrak{F} + AS$.

Come si ricorderà, l'ipotesi del continuo afferma che $2^{\aleph_0} = \aleph_1$ così che – poiché in generale si ha $\aleph_1 \leq 2^{\aleph_0}$ – affinché un modello sia in grado di falsificarla basterà che tra i suoi elementi contenga insiemi X_1 e X_2 le cui cardinalità κ_1 e κ_2 siano tali che

$$\aleph_0 < \kappa_1 < \kappa_2 \leq 2^{\aleph_0}.$$

Il modo più diretto per definire il modello costruito da Cohen è di estendere la nozione di condizione in maniera da permettere che questo ruolo sia svolto non solo da insiemi coerenti di formule atomiche e negazioni di atomiche, ma da elementi *qualunque* di insiemi ordinati. È questo l'approccio indicato da R. Solovay che sarà poi sviluppato in dettaglio da Shoenfield nel 1971 e che è alla base della nozione di *proprietà di forcing* oggi in uso. Fatto questo, la strategia seguita da Cohen ammette una descrizione molto semplice. Consideriamo il modello minimale $\mathfrak{M}$ e siano κ_1 e κ_2 gli insiemi che, in $\mathfrak{M}$, svolgono il ruolo rispettivamente di $\aleph_1$ e $2^{\aleph_0}$. Per definizione, i cardinali sono ordinali iniziali e per l'assolutezza della nozione di ordi-

nale avremo che κ_1 e κ_2 sono ordinali – ordinali numerabili in quanto come sappiamo $\mathfrak{M}$ è un insieme numerabile. Per il resto sicuramente κ_1 e κ_2 non coincideranno con $\aleph_0$ e $2^{\aleph_0}$ nell'universo in cui studiamo $\mathfrak{M}$, ed ω, l'ordinale costituito da tutti gli ordinali finiti, svolgerà il ruolo di $\aleph_0$ tanto in $\mathfrak{M}$ che nell'universo della metateoria.

L'idea di Cohen è ora quella di costruire un'estensione $\mathfrak{N}_2$ di $\mathfrak{M}$ in cui κ_1 diviene più che numerabile e κ_2 ha cardinalità maggiore di κ_1 ma minore di $2^{\aleph_0}$. In questo modo avremo che

$$\aleph_0 < \kappa_1 < \kappa_2 \leq 2^{\aleph_0}$$

e sarà quindi falsificata l'ipotesi del continuo. Tutto sta nell'imporre a κ_1 e κ_2 le proprietà in questione ed è a questo punto che scatta la scelta specifica delle condizioni di *forcing*. Come condizioni prenderemo le parti *finite* del prodotto $\kappa_2 \times \omega$ (vale a dire gli insiemi finiti di coppie di ordinali $< \alpha, n >$ con $\alpha < \kappa_2$ ed n numero naturale). Al solito, considereremo le condizioni come ordinate dall'inclusione e una volta introdotto il linguaggio opportuno con una costante $\bar{a}$ per denotare l'insieme generico che vogliamo aggiungere a $\mathfrak{M}$, definiremo nel modo ovvio la relazione di *forcing* per formule *atomiche*: l'estensione a tutte le formule avverrà nella maniera ormai consueta. Seguendo la falsariga indicata nel caso di $V = L$, otterremo così un'estensione $\mathfrak{N}_2$ di $\mathfrak{M}$ che contiene un insieme generico a ed in cui una formula $\mathscr{A}$ è vera se e solo se è forzata da almeno una condizione nella successione completa $\{P_n\}$ che individua il modello stesso.

Da $\bar{a}$ possiamo ora costruire una funzione $F{:}\kappa_2 \rightarrow \mathscr{P}(\omega)$ ponendo

$$F(\alpha) = \{n \in \omega \mid < \alpha, n > \in \bar{a}\}$$

ed è facile verificare che F sarà iniettiva, così che $\kappa_2 \leq 2^{\aleph_0}$. Ragionando ora per assurdo e sfruttando il fatto che $\aleph_2$ è generico, potremo provare che ω, κ_1 e κ_2 sono cardinali *distinti* e quindi che *IC* è *falsa* in $\mathfrak{N}_2$. Metodi analoghi permettono di falsificare anche *IGC* e più in generale di alterare in modo radicale le cardinalità passando da un modello a sue estensioni generiche. L'idea di fondo è sempre quella di «codificare» attraverso i generici (prendendo quindi come condizioni funzioni parziali di dominio finito o insiemi di coppie la cui seconda coordinata è naturale) funzioni che

hanno come dominio ω e stabiliscono una biiezione con cardinali che nel modello di partenza sono più grandi. In questo modo nel nuovo modello è possibile «*contrarre*» («collapse») i cardinali e costringerli ad assumere valori più bassi. I primi esempi di questa tecnica furono dati da Cohen stesso, ma furono soprattutto Levy e Feferman a sviluppare in dettaglio il metodo.

Più che sui risultati specifici che a partire dalle prime comunicazioni di Cohen si susseguirono – all'inizio con ritmo febbrile – ad opera di diversi ricercatori, è sullo sviluppo del metodo del *forcing* in se stesso che conviene soffermarci ancora, in quanto esso non solo aprì orizzonti completamente nuovi allo studio della teoria degli insiemi e delle sue applicazioni ma – sul piano più generalmente logico – costituì il punto d'incontro di temi e problematiche sino ad allora separati.

Fu solo lentamente che la struttura logica del metodo emerse in modo netto. All'inizio, come ricordato, Cohen definiva la relazione $P \Vdash \mathscr{A}$ solo per formule in forma prenessa e l'attenzione era rivolta al comportamento dei quantificatori (in particolare $\forall$) più che alla negazione. Fu Scott a porre in rilievo il ruolo centrale di $\neg$ e a sottolineare un fatto che nella presentazione originaria di Cohen non emergeva: lo stretto parallelismo tra le clausole per $\wedge$, $\vee$, $\neg$, e $\exists$ nella definizione di *forcing* e quelle date da Kripke per la verità intuizionista. Di fatto l'analogia aveva un significato profondo che diversi ricercatori non tardarono a rendere esplicito.

In un suo lavoro del 1964, (quindi *prima* della pubblicazione dei lavori di Kripke al riguardo) Grzegorczyk mostrava come – se ci si limitava a formule prive di quantificatori – gli enunciati forzati da ogni condizione in ogni insieme ordinato $< I, \leq >$ coincidevano con i teoremi proposizionali intuizionisti. Questo collegamento sembrava estendersi anche ad alcune ricerche di Kreisel risalenti al 1959, in cui la logica intuizionista veniva reinterpretata come la logica riguardante totalità in potenza e non in atto, come nel caso della logica classica, e quindi come la logica naturale per gli oggetti che vengono costruiti passo per passo. Del resto – dopo che la tecnica di Cohen aveva mostrato le sue potenzialità entro la teoria degli insiemi – non furono pochi i logici che riconobbero in essa spunti che provenivano da contesti differenti; innegabili erano i contatti – oltre che con la logica intuizionista – con i metodi di

priorità utilizzati nello studio dei gradi, in particolare nella costruzione passo passo di gradi fra loro incomparabili (come nella dimostrazione di Post e Kleene del 1954).

Quest'aria di famiglia nasceva dalla somiglianza di problemi e imponeva un'analisi più ravvicinata della struttura logica del metodo. Così, dopo che Scott aveva mostrato il ruolo centrale della clausola sulla negazione nella definizione di *forcing*, Feferman isolò il concetto di *forcing debole* che da diversi punti di vista è più maneggevole di quello di Cohen. Una condizione P forza debolmente $\mathscr{A}$ ($P \Vdash^{*} \mathscr{A}$) se $P \Vdash \neg\neg \mathscr{A}$. Come nel caso intuizionista, il passaggio alla doppia negazione stabilisce un contatto più stretto con la logica e la verità classica. Così si ha che il *forcing debole* rispetta l'equivalenza logica classica e ammette una caratterizzazione cui abbiamo già accennato e che è fondamentale: $P \Vdash^{*} \mathscr{A}$ se e solo se $\mathscr{A}$ è vera in *ogni* modello generico indotto da successioni complete che contengono P.

È questa la chiave di volta del collegamento fra genericità e verità classica e se si vede la cosa in termini di strutture di Kripke si percepisce subito l'analogia con i teoremi di Glivenko che collegano dimostrabilità intuizionista e dimostrabilità classica via passaggio alla doppia negazione: la costruzione di successioni complete entro le strutture di Kripke offre un modo per ridimostrare e generalizzare i risultati di Glivenko. Il legame sarebbe emerso in tutta la sua forza quando nel 1969 Melvin C. Fitting, nel suo *Intuitionistic Logic Model Theory and Forcing* (*Logica intuizionista, teoria dei modelli e forcing*) avrebbe dato una versione dei risultati di indipendenza di Cohen partendo da gerarchie di strutture di Kripke costruite in analogia con quella dei costruibili.

Utilizzando strutture di questo tipo Fitting era ora in grado di provare la non dimostrabilità, sulla base della logica *intuizionista*, dei vari principi in esame (*AS*, *IGC*, ecc.) dagli assiomi di $\mathfrak{Z}\mathfrak{F}$. A questo punto scattava una forma del teorema di Glivenko che permetteva di provare la non dimostrabilità – anche dal punto di vista *classico* – di tali principi. Che vantaggi offriva questa procedura? Rispetto alla strategia originaria il vantaggio era che in questo modo non era necessario costruire modelli *classici per* falsificare *AS* o *IGC* o $V = L$: bastava limitarsi a quelli intuizionisti. Al posto delle estensioni generiche e delle successioni complete si faceva ricorso al teorema di Glivenko che collega dimostrabilità classica con dimostrabilità intuizionista.

Questo poneva in luce due fatti che si sarebbero rivelati estremamente importanti. Il primo riguardava le ipotesi necessarie per applicare il metodo di costruzione dei generici. Cohen, come si ricorderà, partiva dal modello minimale di cui provava l'esistenza assumendo per ipotesi l'esistenza di un modello standard numerabile. Di fatto si trattava di un artificio stilistico per semplificare l'esposizione, in quanto Cohen sapeva benissimo come ottenere l'esistenza del modello desiderato partendo dalla semplice coerenza di $\mathfrak{ZF}$. Questo però non mancò di sollevare perplessità sia dal punto di vista puramente metodologico che da quello pratico: che limiti aveva il metodo di Cohen? Sotto che ipotesi era utilizzabile? La risposta che non tardò a venire da più parti fu che la costruzione esplicita di modelli non era necessaria e tanto meno che era inevitabile il ricorso a modelli numerabili e a modelli standard. Come Fitting mostrava la dispensabilità dei modelli classici, così, a partire dal 1965, Petr Vopĕnka e i suoi allievi (la cosiddetta scuola di Praga) ponevano in luce altre soluzioni che eliminavano questa necessità. Sempre nei primi anni sessanta Robert Solovay indicava come il *forcing* si potesse applicare anche partendo da modelli non standard e tutto ciò non faceva che sottolineare come i nuovi metodi avessero una portata estremamente ampia.

Rimaneva il secondo fatto che l'approccio di Fitting illustra in modo eloquente: dimostrare l'indipendenza di un enunciato $\mathscr{A}$ da una teoria $\mathfrak{T}$ dal punto di vista della logica classica non significa necessariamente che ci si deve limitare a interpretazioni classiche. Come posto in luce da Fitting, si può benissimo ricorrere ad una logica non classica come quella intuizionista ed utilizzare poi teoremi di collegamento tipo quello di Glivenko. In quest'ottica il forcing è la logica dei modelli che usiamo per falsificare le varie ipotesi in esame e le successioni complete hanno il *solo* ruolo di permetterci di passare dalla non dimostrabilità rispetto a questi modelli a quella classica. Ciò dà un senso ben preciso alle tecniche impiegate: usare modelli intuizionisti (i modelli la cui logica è data dalla relazione di *forcing*) è utile in quanto essi includono propriamente quelli classici e risultano intuitivamente duttili sul piano euristico. È quindi più facile trovare contromodelli che non riusciremmo a costruire rimanendo entro l'universo classico. In altre parole, una volta visti in questa ottica, i metodi introdotti da Co-

hen costituiscono un invito – anche quando ci si occupa di teorie *classiche* – ad allargare l'universo delle interpretazioni in modo da disporre di più metodi di costruzione.

La via seguita da Fitting rendeva esplicito il legame tra logica intuizionista e *forcing* ma più o meno nello stesso giro d'anni si andava precisando un altro modo di presentare il *forcing* che, se non costituiva un distacco dalla logica classica, comportava comunque una profonda modificazione del concetto di interpretazione. Si tratta dei *modelli booleani* che da prospettive diverse e pressoché simultaneamente furono introdotti da Vopěnka e la sua scuola in Cecoslovacchia e da Solovay e Scott negli Stati Uniti. Di fatto furono soprattutto le note di Scott – che circolarono dattiloscritte per molto tempo – a diffondere la conoscenza del metodo mentre i lavori della scuola praghese rimasero nell'ombra anche se importanti contributi nati in questo contesto teorico (ad esempio i teoremi di A. Sochor e T. Jech che mostravano come fosse possibile immergere sotto opportune condizioni modelli di permutazioni con atomi in modelli booleani) furono presto riformulati in termini della teoria di Scott-Solovay.

L'idea dei modelli booleani nasceva da due osservazioni fondamentali riguardanti il *forcing*. La prima – cui abbiamo già accennato – è che il concetto di *forcing* in sé e per sé non riguarda esclusivamente condizioni date da formule. Possiamo applicare il metodo a partire da insiemi ordinati $< I, \leq >$ *qualsiasi*, una volta che si stipuli caso per caso quali formule atomiche dovranno essere forzate dalle condizioni: le clausole sui connettivi e i quantificatori fanno riferimento alla pura struttura d'ordine dell'insieme. In questo modo è possibile (l'abbiamo visto, ad esempio, nel caso di *IGC*) trattare in modo uniforme costruzioni in cui le condizioni sono oggetti di tipo diverso: funzioni parziali, insiemi perfetti, boreliani di misura positiva, ecc. Furono proprio questi esempi a sottolineare l'importanza di un approccio più astratto.

In questa prospettiva, il *forcing* si applica a partire da una *nozione di forcing* $< I, \leq >$ che è un insieme ordinato che appartiene al nostro *modello di partenza* $\mathfrak{M}$. Su $\mathfrak{M}$ si assume solo che sia un insieme transitivo e di fatto entrambe le ipotesi potrebbero essere indebolite: l'essenziale è che in ogni caso cade la vincolante sulla numerabilità. Il ruolo delle successioni complete viene preso ora dai

filtri generici su $\mathfrak{M}$, introdotti da Levy nel 1964, in cui per ragioni tecniche l'ordine viene invertito rispetto a quanto faceva Cohen e intuitivamente $x \leq y$ significa che x ha più informazioni di y. Si dice che un sottoinsieme G di I è un *filtro generico* su $\mathfrak{M}$ se:

a) Se $x, y \in G$, esiste uno $z \in G$ tale che $x \geq z$ e $y \geq z$

b) Se $x \in G$ e $y \in I$, allora $y \in G$ nel caso $y \geq x$

c) G ha intersezione non vuota con ogni sottoinsieme *denso* $X \subseteq I$ appartenente a $\mathfrak{M}$.

La nozione di densità è essenziale per capire in che senso si può lasciar cadere la condizione di numerabilità di $\mathfrak{M}$. $X \subseteq I$ è *denso* se nel caso $x \in X$ e $x \geq y$ anche $y \in X$ e se per ogni $x \in I$ esiste un $y \in X$ per cui $x \geqslant y$. Per definizione, se $\mathscr{A}$ è una formula di un qualche linguaggio ramificato come quelli utilizzati da Cohen, è immediato verificare, se invertiamo l'ordine, che l'insieme

$$|\mathscr{A}| = \{p \in I \mid p \Vdash \mathscr{A}\}$$

è denso e per la definibilità della relazione di *forcing* (fissata la condizione) esso apparterrà a $\mathfrak{M}$. Ne scende che per dimostrare l'esistenza di una successione completa di condizioni basterà provare l'esistenza di un filtro generico su $\mathfrak{M}$ e che per questo non occorre postulare che $\mathfrak{M}$ sia numerabile: basta che numerabile sia la famiglia di sottoinsiemi di I che appartengono a $\mathfrak{M}$. Questo è sufficiente per applicare la dimostrazione di Cohen per provare l'esistenza di filtri generici ed è questa la condizione che in generale sottostà all'applicazione del *forcing*.

Di altre ipotesi più forti avremo occasione di parlare nel prossimo capitolo; per ora ci basti osservare che in questa impostazione astratta, fissato un filtro generico G su $\mathfrak{M}$, l'estensione generica $\mathfrak{M}[G]$ che esso individua (nel modo che abbiamo visto parlando delle dimostrazioni di indipendenza di Cohen) è univocamente caratterizzata dal fatto di essere un modello transitivo di $\mathfrak{ZF}$ che estende $\mathfrak{M}$, contiene tra i suoi elementi G ed è il *più piccolo* modello transitivo che fa questo, nel senso che se $\mathfrak{N} \supseteq \mathfrak{M}$ e $G \in \mathfrak{N}$, avremo che $\mathfrak{M}[G] \subseteq \mathfrak{N}$ nel caso $\mathfrak{N}$ sia modello di $\mathfrak{ZF}$. Altro fatto importan-

te è che se $\mathfrak{M}$ soddisfa AS, anche $\mathfrak{M}[G]$ lo farà e che – abbiamo visto il ruolo essenziale di questa circostanza provando l'indipendenza di $V = L$ – $\mathfrak{M}[G]$ e $\mathfrak{M}$ hanno *gli stessi ordinali.*

Poste le cose in questo modo, è chiaro come siano la struttura d'ordine e il concetto di densità a reggere le fila dell'intera costruzione. Ma c'è un altro fatto da osservare e questo è il secondo punto fondamentale che Solovay mise in luce nel 1965. Fissata una formula $\mathscr{A}$, con $|\mathscr{A}|$ abbiamo indicato l'insieme delle condizioni che la forzano. Si tratta di un sottoinsieme di I, quindi di un elemento di $\mathscr{P}(I)$. Associamo ora ad $\mathscr{A}$ la coppia

$$\|\mathscr{A}\| = < |\mathscr{A}|, |\neg \mathscr{A}| >$$

delle condizioni che forzano $\mathscr{A}$ e di quelle che forzano $\neg \mathscr{A}$. Si può dimostrare che tutti i valori $\|\mathscr{A}\|$, dove $\mathscr{A}$ è una formula, risultano elementi di un'algebra di Boole completa. Il fatto importante è che si verificano le condizioni seguenti che mostrano come la definizione di *forcing* si lascia analizzare in termini di un'assegnazione alle formule di valori di verità nell'algebra di Boole $\mathfrak{B}$. Abbiamo infatti

$$\|\neg \mathscr{A}\| = \sim \|\mathscr{A}\|$$
$$\|\mathscr{A} \wedge \mathscr{B}\| = \|\mathscr{A}\| \cap \|\mathscr{B}\|$$
$$\|\mathscr{A} \vee \mathscr{B}\| = \|\mathscr{A}\| \cup \|\mathscr{B}\|$$
$$\|\forall x \mathscr{A}\| = \cup_{a \in S} \|\mathscr{A}[\bar{a}]\|$$
$$\|\exists x \mathscr{A}\| = \cup_{a \in S} \|\mathscr{A}[\bar{a}]\|.$$

La teoria dei modelli booleani nasce quando si verifica, come ben presto fecero Scott e Solovay, che si può fare il passo inverso e che è possibile costruire analoghi dei modelli generici a partire da assegnazioni di valori di verità entro algebre di Boole complete. Il vantaggio non è solo quello di permettere di fare a meno di modelli numerabili e di assioma di costruibilità, ma è soprattutto di carattere metodologico: la costruzione dei modelli booleani è estremamente più semplice sul piano matematico dell'uso del *forcing*, ed è anche più diretta. Meno chiara è la dipendenza tra proprietà delle algebre di Boole che di volta in volta si usano e i modelli costruiti. In molti casi per fare questa scelta si ricorre ad interpreta-

zioni dei modelli in termini di *forcing*. Ciò che è indubitabile in ogni modo è la limpidezza intuitiva: i modelli booleani danno un'immagine diretta del modo in cui questi nuovi universi generalizzano quelli classici.

L'idea di fondo è di porre in primo piano le *funzioni caratteristiche* al posto degli insiemi. Dato l'insieme X sappiamo che ogni $Y \subseteq X$ è univocamente determinato dalla funzione $f_Y : X \to \mathbf{2}$ (dove $\mathbf{2}$ è l'algebra semplice con i soli elementi 0 e 1) definita ponendo

$$f_Y(a) = \begin{cases} 1 & \text{se } a \in Y \\ 0 & \text{altrimenti.} \end{cases}$$

In questa maniera l'appartenenza o meno di a al sottoinsieme Y si valuta in modo netto: o $a \in Y$ (e allora $f_Y(a) = 1$) o $a \notin Y$ (e allora $f_Y(a) = 0$): non ci sono casi intermedi, appartenenze parziali o sfumate.

Valutare in un'algebra di Boole $\mathfrak{B}$ significa ammettere questa eventualità ed identificare i sottoinsiemi di X con funzioni $f: X \to \mathfrak{B}$. A questo punto, come gli universi classici per $\mathfrak{ZF}$ si ottengono nella gerarchia di von Neumann iterando l'operazione di passaggio all'insieme delle parti, è naturale costruire i nostri universi booleani definendo una gerarchia $\{V_\alpha(\mathfrak{B})\}$ per ogni $\alpha \in Ord$. Per far questo occorre però modificare l'idea originaria di identificare gli insiemi con funzioni $f: X \to \mathfrak{B}$ in quanto, al posto di X, dovremo avere funzioni con valori in $\mathfrak{B}$. La soluzione cui si arriva è la seguente che – come si può provare – è l'unica che soddisfi plausibili condizioni:

$$V_0 = \emptyset$$

Per ogni $\alpha \in Ord$

$$V_{\alpha+1}^{(\mathfrak{B})} = \{x \in \mathrm{Fun}\,(x) \mid \mathrm{Cod}\,(x) \subseteq B \text{ e } \mathrm{Dom}(x) \subseteq V_\alpha^{(\mathfrak{B})}\}.$$

In altre parole, $V_{\alpha+1}^{(\mathfrak{B})}$ è la collezione delle funzioni x il cui codominio (insieme dei valori) è incluso nell'algebra $\mathfrak{B}$ e i cui argomenti sono elementi di $V_\alpha^{(\mathfrak{B})}$. Per ordinali limite, come nel caso di von Neumann, si farà la riunione sui livelli precedenti e si giunge così alla definizione dell'universo

$$V^{(\mathfrak{B})} = \bigcup_{\alpha \in Ord} V_\alpha^{(\mathfrak{B})}.$$

È facile provare non solo che gli universi classici si possono vedere come universi booleani in cui $\mathfrak{B} = \mathbf{2}$, ma che essendo $\mathbf{2}$ sottoalgebra di *ogni* algebra di Boole, potremo vedere V come inclusa in $V^{(\mathfrak{B})}$ ed in questo senso i modelli booleani ci permettono di *estendere* i modelli classici. L'arricchimento che otteniamo passando a questi universi è dato dalla maggior articolazione interna dei loro elementi, dalla complicazione che essi possono presentare non essendo più limitati alla rigida dicotomia Vero/Falso.

Rimane da definire il concetto di appartenenza tra gli insiemi booleani che abbiamo così introdotto e questo – se vogliamo che valga un principio d'estensionalità – comporta una definizione intrecciata con quella di uguaglianza. Come nel caso del *forcing* procederemo per induzione sul rango degli insiemi (cioè sul minimo livello della gerarchia cui essi appartengono) ed avremo:

$$\|x \in y\| = \bigcup_{z \in Dom(y)} \{\|y(z) \cap z = x\|\}$$

$$\|x = y\| = \bigcap_{z \in Dom(x)} \{x(z) \rightarrow \|z \in y\|\} \bigcap_{w \in Dom(y)} \{y(w) \rightarrow \|w \in x\|\}.$$

Possiamo vedere queste condizioni come le clausole per le formule atomiche della definizione induttiva di una funzione

$$\| \quad \| : L \rightarrow \mathfrak{B}$$

che associa un valore di verità in $\mathfrak{B}$ ad *ogni* formula $\mathscr{A}$ del linguaggio L. L'estensione si otterrà usando proprio quelle clausole su $\wedge$, $\vee$, $\neg$, $\forall$, $\exists$, che abbiamo visto sopra parlando del passaggio dal *forcing* ai modelli booleani.

Ciò che rende possibile l'uso di questi modelli nello studio di $\mathfrak{ZF}$ dal punto di vista classico è che – malgrado la loro maggiore articolazione interna rispetto ai modelli standard – i modelli booleani *sono* oggetti classici, in un senso preciso, in quanto soddisfano la logica classica e rendono veri tanto gli assiomi di $\mathfrak{ZF}$ che AS (una volta che si parta da modelli classici di questo assioma). Si può provare infatti che se $\mathscr{A}$ è una legge logica classica, un assioma di $\mathfrak{ZF}$ o AS avremo che

$$\|\mathscr{A}\| = 1.$$

Tutto ciò si può dimostrare all'interno di $\mathfrak{ZF}$ in quanto, come nel *forcing*, la classe $V^{(\mathfrak{B})}$ è definibile (con un parametro $\mathfrak{B}$ per indicare l'algebra) proprio come l'universo V o quello L dei costruibili. La stessa procedura ci permette di costruire un modello $\mathfrak{M}^{(\mathfrak{B})}$ a partire da un modello standard $\mathfrak{M}$ di $\mathfrak{ZF}$, ma l'importante in generale è che – diversamente dal metodo del *forcing* quale presentato originariamente da Cohen – qui non è necessario partire da modelli $\mathfrak{M}$ particolari, a meno che questo non sia richiesto da specifiche considerazioni. Per dimostrare che un dato enunciato $\mathscr{A}$ non è teorema di $\mathfrak{ZF}$, ci basterà trovare un modello booleano $\mathfrak{M}^{(\mathfrak{B})}$ in cui si ha $\|\mathscr{A}\| \neq 1$: da questo – in quanto la logica classica è valida per i modelli booleani – scenderà che $\mathscr{A}$ *non* è dimostrabile in $\mathfrak{ZF}$.

Come il lettore ricorderà, questo fatto, come la nozione generale di *struttura booleana* non erano nuovi nella letteratura ma erano stati oggetto di studio sistematico da parte di Mostowski e della sua scuola già negli anni cinquanta. Un'esposizione estremamente sofisticata dal punto di vista algebrico si trovava nel classico *La matematica della metamatematica* di Rasiowa e Sikorski da noi ricordato nel capitolo II. I due autori mostravano la validità e la completezza della logica classica rispetto alle strutture booleane, come facevano pure per la logica intuizionista rispetto ad analoghe strutture con valori in algebre di Heyting. Non usavano però in concreto questi metodi per studiare e costruire modelli di teorie specifiche e non affrontavano il passaggio decisivo al secondo ordine o alla teoria dei tipi che è essenziale per interpretare la teoria degli insiemi.

Idee al riguardo erano state espresse nel 1951 da Alonzo Church – seguendo un suggerimento del suo allievo Lagerström – ma sostanzialmente non c'era nulla di concreto. Questo spiega il successo a sorpresa che i modelli booleani ebbero presso i matematici, soprattutto presso quelli che preferivano un approccio ai problemi di indipendenza che riducesse al minimo l'uso della logica. In questo senso un ruolo notevole nella divulgazione dei risultati di indipendenza di Cohen, Solovay, Easton, ecc., di alcuni dei quali parleremo più avanti, ebbe il libro di J.B. Rosser *Simplified independence proofs: Boolean valued models of set theory* (*Dimostrazioni semplificate d'indipendenza: i modelli booleani per la teoria degli insiemi*)

pubblicato nel 1969, in cui i modelli booleani venivano utilizzati come strumento principale. Rosser era stato negli anni cinquanta l'apostolo delle logiche polivalenti presso i matematici e non mancava nell'introduzione di sottolineare il significato di questa ripresa di prospettive antagoniste alla bivalenza.

C'era però un ulteriore collegamento con nozioni ormai classiche nella logica che rendeva ancora più plausibile e motivato il passaggio dai modelli classici a quelli booleani e questo era fornito dal concetto di ultraprodotto che – è un fatto spesso dimenticato – era stato introdotto nel 1955 da Jerzy Łos proprio in una prospettiva polivalente, anche se era stato ben presto riassorbito dai lavori di Frayne, Morel e Scott nel contesto della teoria dei prodotti ridotti di algebre. Gli ultraprodotti – come fece vedere Scott in una sua breve nota del 1969 – si possono vedere come ottenuti in due passi: un primo passo ci permette di definire, a partire dalle strutture fattori $\{\mathfrak{A}_i\}_{i\in I}$ il dominio $\Pi_{i\in I} A_i$ di una struttura booleana $\mathfrak{A}$ in cui ogni costante relazionale n-aria R è interpretata in modo che

$$\|R(f_1,\dots,f_n)\| = \{i \in I \mid \mathfrak{A}_i \models R[f_1(i),\dots,f_n(i)]\} = F_R(f_1,\dots,f_n)\}$$

$$\|f = g\| = \{i \in I \mid f(i) = g(i)\}$$

dove f, g sono elementi del dominio di $\mathfrak{A}$ e quindi funzioni.

Un secondo passo in cui da questa struttura booleana, quozientando, passiamo ad una struttura classica usando un opportuno ultrafiltro F nell'algebra di Boole $\mathscr{P}(I)$ delle parti di I e otteniamo l'ultraprodotto $\Pi_{i\in I}\mathfrak{A}_i/F$. Il teorema di Łos che sta alla base della teoria degli ultraprodotti e vincola la verità di una formula nell'ultraprodotto alla sua verità in un insieme «grande» di strutture («grande» significa che la famiglia degli indici appartiene all'ultrafiltro) viene ora a significare che c'è uno stretto rapporto tra la verità booleana in $\mathfrak{A}$ e quella classica in $\Pi_{i\in I}\mathfrak{A}_i/F$ nel senso che, poiché $I \in F$ avremo:

$$\mathfrak{A} \models \mathscr{A} \Rightarrow \Pi_{i\in I}\mathfrak{A}_i/F \models \mathscr{A}.$$

Le strutture booleane costituiscono una generalizzazione del primo passo della costruzione degli ultraprodotti e Scott proprio utilizzando un modello booleano particolarmente intuitivo, i cui elementi

sono funzioni di Borel, dava nel 1967 quella che è forse a tutt'oggi la dimostrazione più semplice e diretta della indipendenza di *IC*.

Il ricorso a ultrafiltri per passare da strutture booleane a strutture classiche è però anche la chiave di volta per afferrare il collegamento tra modelli booleani e modelli generici. Supponiamo infatti che $\mathfrak{M}$ sia un modello transitivo di $\mathfrak{ZF}$ e che $\mathfrak{B} \in \mathfrak{M}$ sia un'algebra di Boole completa *in* $\mathfrak{M}$ (che contiene cioè il sup e l'inf di quei sottoinsiemi di $\mathfrak{B}$ che sono elementi di $\mathfrak{M}$). Se $\mathfrak{M} \cap \mathscr{P}(\mathfrak{B})$ è numerabile, il lemma di Rasiowa e Sikorski ricordato nel capitolo II ci garantisce che per ogni $b \in \mathfrak{B}$ esiste un filtro G su $\mathfrak{B}$ che contiene b e ha intersezione non vuota con tutti i sottoinsiemi densi di $\mathfrak{B}$ che appartengono a $\mathfrak{M}$. In altre parole (ricordando che le definizioni date parlando di nozione di *forcing* si estendono ad algebre di Boole in quanto queste sono insiemi ordinati) avremo un filtro generico G su $\mathfrak{M}$ che – è immediato – sarà pure un ultrafiltro. Si noti che in generale (come per i generici nella costruzione di Cohen) G *non* sarà un elemento di $\mathfrak{M}$. Ciò non toglie che ora possiamo quozientare $\mathfrak{M}^{(\mathfrak{B})}$ come nella costruzione degli ultraprodotti analizzata da Scott ed ottenere un modello classico $\mathfrak{M}^*[G]$ che rende vere tutte le formule che hanno valore 1 in $\mathfrak{M}^{(\mathfrak{B})}$.

$\mathfrak{M}^*[G]$ non sarà in generale un modello standard ma soddisferà l'estensionalità e sarà transitivo. Applicando l'isomorfismo di contrazione di Mostowski otterremo quindi un modello standard $\mathfrak{M}[G]$ che rende vere le formule vere in $\mathfrak{M}^{(\mathfrak{B})}$, contiene G ed estende $\mathfrak{M}$, soddisfacendo la caratterizzazione delle estensioni generiche $\mathfrak{M}[G]$ che abbiamo date a suo tempo parlando del concetto generale di nozione di *forcing*. È quindi possibile costruire le estensioni generiche partendo dai modelli booleani e si noti che l'ipotesi che abbiamo utilizzato per applicare il lemma di Rasiowa-Sikorski riguardante la numerabilità di $\mathfrak{M} \cap \mathscr{P}(\mathfrak{B})$ è sostanzialmente la stessa che in termini di nozioni di *forcing* vincola l'esistenza di filtri generici alla numerabilità di $\mathfrak{M} \cap \mathscr{P}(I)$.

Da questo collegamento scende un'ultima e decisiva semplificazione delle tecniche di Cohen che fu introdotta in maniera esplicita da Shoenfield nel 1971, ed è l'abbandono dei linguaggi ramificati. La costruzione di estensioni generiche via modelli booleani non richiede procedure passo passo e non fa appello a linguaggi che non siano quelli elementari così che è questa ben più semplice versione non ramificata che è oggi in uso.

Sul piano tecnico, d'altra parte, non tardarono a presentarsi altri raffinamenti ed approfondimenti dei metodi tanto del *forcing* quanto dei modelli booleani: tra questi, lo studio dei rapporti tra contrazione dei cardinali e condizioni di catena, la questione della iterabilità del *forcing* per costruire successivamente più estensioni generiche, ecc. Non mancarono anche – fin dal loro primo apparire – le applicazioni dei modelli booleani che andavano nel senso inverso e permettevano di ottenere risultati sulle algebre di Boole a partire da modelli booleani. È questo il caso dell'estensione del teorema di Gaifman-Hales data prima da Solovay nel 1965 e poi da Kripke nel 1967 sull'immergibilità completa di algebre di Boole di cardinalità arbitrariamente alta. Lo studio dello sfondo algebrico e topologico di *forcing* e modelli booleani è stato un tema particolarmente intenso d'indagine a partire dalla metà degli anni sessanta, che ha gettato luce inaspettata non solo sui modelli, ma sulla stessa topologia generale e sulle algebre di Boole. Importanti in questo contesto sono i lavori di Takeuti sulla versione topologica del *forcing*, come quelli di Mostowski, G. Reyes e altri.

Collegata alla rinascita di interesse per le strutture booleane è anche tutta una serie di lavori che riguardano non tanto la teoria degli insiemi quanto quella dei modelli. Un posto a parte ha la creazione del *forcing* finito ed infinito di Robinson, che portò ad importanti sviluppi nella *model-theoretic Algebra* e di cui ci occuperemo più avanti. In un'altra direzione si situano invece i lavori di D. Scott e P. Krauss sui modelli probabilistici, che costituiscono un ponte tra logica induttiva nel senso di Carnap, studio delle strutture booleane (qui si considerano algebre di misura, come nella teoria della probabilità) e teoria generale dei modelli. I lavori di Scott e Krauss riprendono precedenti tentativi di Łos, Fenstad e Gaifman, ma è l'esplicito riferimento ai modelli booleani che dà un sapore nuovo ai loro contributi.

Gli esempi che abbiamo dato sino ad ora illustrano d'altra parte solo *alcuni* degli sviluppi che la problematica connessa al *forcing* e più in generale ai modelli booleani ha realizzato a partire dai primi anni sessanta e la situazione è oggi notevolmente più articolata. Un tipo di indagini che ha preso le mosse dai problemi di indipendenza della teoria degli insiemi e su cui non ci è possibile soffermarci in modo adeguato malgrado il loro estremo interesse è costituito dagli sviluppi della cosiddetta *teoria dei semiinsiemi* (semi-

sets) creata da Vopěnka e dal suo allievo P. Hajek nel 1967 e che è stata al centro dell'attività della scuola di Praga. Abbiamo già ricordato come l'approccio di Vopěnka ai problemi di indipendenza fosse sostanzialmente equivalente sul piano matematico alla teoria dei modelli booleani sviluppata da Solovay e Scott; quello che però va sottolineato è la diversa prospettiva in cui Vopěnka e Hajek si pongono dal punto di vista metodologico. Diversamente da Cohen che costruiva le estensioni generiche partendo da modelli che già assumeva di avere e sui quali imponeva ipotesi restrittive come la numerabilità, l'approccio di Vopěnka e Hajek è di tipo sintattico. Dare un modello per una teoria $\mathfrak{T}$ significa definire una *trasformazione sintattica* che a formule della teoria $\mathfrak{T}$ associa formule di una teoria $\mathfrak{T}'$ rispettando la proprietà di essere teorema, nel senso che se $\mathfrak{T} \vdash \mathscr{A}$, dovremo avere $\mathfrak{T}' \vdash \mathscr{A}^{\mathfrak{M}}$, dove $\mathscr{A}^{\mathfrak{M}}$ è la formula che il modello $\mathfrak{M}$ associa ad $\mathscr{A}$.

Si tratta di una nozione che generalizza quella di modello interno di Gödel e quella di interpretazione di Tarski che – come abbiamo visto sopra parlando della definibilità della relazione $\mathfrak{M}^{(\mathfrak{B})} \models \mathscr{A}$ – comprende come caso particolare anche lo studio dei modelli booleani e quindi del *forcing*. Vopěnka e Hajeck sottolineano come il contesto naturale per lo sviluppo dello studio dei modelli sintattici sia la teoria delle categorie e non è difficile trovare forti analogie tra i modelli nel loro senso e la semantica funtoriale introdotta da William Lawvere proprio nello stesso giro d'anni.

Il punto che vogliamo sottolineare in questo contesto è però un altro. Vopěnka e Hajek partono da una teoria delle classi $\mathfrak{TC}$ formulata in un linguaggio con $\in$ e S (il predicato «essere un insieme») che sostanzialmente coincide con $\mathfrak{NBG}$ senza gli assiomi dei gruppi A, B, C, e considerano due estensioni: la teoria $\mathfrak{TS}$ degli insiemi (che è appunto equivalente a $\mathfrak{NBG}$) e una teoria intermedia tra $\mathfrak{TC}$ e $\mathfrak{TS}$, la teoria $\mathfrak{TSS}$ dei *semisets*. I semisets sono già definibili in $\mathfrak{TC}$ e sono le sottoclassi d'insiemi o – se si preferisce – classi incluse in insiemi. In $\mathfrak{TS}$, in vista del principio di isolamento, non esistono semisets propri in quanto ogni sottoclasse di un insieme è un insieme. I semisets sono intermedi fra le classi (ogni semiset è una classe) e gli insiemi (ogni insieme è un semiset) e dal punto di vista metodologico il loro interesse sta nel fatto che ci è possibile costruire *modelli sintattici* di $\mathfrak{TS}$ in $\mathfrak{TSS}$ o omettendo clas-

si proprie o trasformando semisets in insiemi. In questo senso la teoria dei semisets è la chiave di volta per analizzare sistematicamente i modelli di $\mathfrak{TS}$ in un universo più ampio, in cui si ammette l'esistenza di oggetti più generali degli insiemi.

Non deve stupire quindi che – a partire dalla metà degli anni settanta – sempre per opera di Vopěnka – da essa sia nata una vera e propria *teoria alternativa degli insiemi* ($\mathfrak{AST}$) in cui è possibile sviluppare parte della matematica senza assumere – come nella teoria cantoriana – l'infinito attuale. In questo contesto, *finite* sono le classi per cui ogni sottoclasse è un insieme così che infiniti sono gli insiemi che hanno semisets *propri* tra le loro sottoclassi. Se si considera il ruolo della definibilità nel trasformare ogni semiset in un insieme, si ha così che infiniti sono gli insiemi in cui alcune proprietà non sono in grado di determinare nettamente elementi che siano insiemi e non semisets. Le ricerche in questa direzione sono oggi attive non solo in Cecoslovacchia e hanno mostrato come $\mathfrak{AST}$ presenti notevoli analogie con l'Analisi non standard di Robinson ed in generale si possa vedere come un contesto naturale in cui sviluppare approcci non standard alla matematica nel suo complesso.

C.C. Chang e H.J. Keisler, *Teoria dei modelli*, Boringhieri, Torino 1980 [edizione originale 1973].

A. Robinson, *Introduzione alla teoria dei modelli e alla metamatematica dell'algebra*, Boringhieri, Torino 1974 [edizione originale 1963].

A. Robinson, *Selected Papers*, 3 volumi, North Holland, Amsterdam 1979.

AA.VV., *Studies in model theory*, a cura di Michael Morley, MAA, Princeton 1973.

G. Sacks, *Saturated model theory*, Benjamin, Reding (Mass.) 1972.

C. Karp, *Languages with Expressions of Infinite Lenght*, North Holland, Amsterdam 1963.

AA.VV., *The Syntax and Semantics of Infinitary, Languages*, a cura di Jon Barwise, Springer Verlag, Berlino 1968.

H.J. Keisler, *Model Theory for Infinitary Languages*, North Holland, Amsterdam 1971.

M.A. Dickman, *Large Infinitary Languages*, North Holland, Amsterdam 1975.

R. Smullyan, *First order Logic*, Springer Verlag, Berlino 1968.

AA.VV., *Intuitionism and Proof Theory*, a cura di Akiko Kino, John Myhill, Richard Eugene Vesley, North Holland, Amsterdam 1970.

AA.VV., *Proceedings of the second Scandinavian Logic Symposium*, a cura di Jens Erik Fenstad, North Holland, Amsterdam 1971.

K. Schütte, *Proof Theory*, Springer Verlag, Berlino 1977 [edizione originale 1960].

J.Y. Girard, *Proof Theory and Logical Complexity*, I volume, Bibliopolis, Napoli 1987.

S.C. Kleene e R.E. Vesley, *The Foundations of Intuitionistic Mathematics*, North Holland, Amsterdam 1965.

AA.VV., *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, a cura di Anne Sjèpr Troelstra, Springer Verlag, Berlino 1973.

M. Dummett, *Elements of Intuitionism*, Clarendon Press, Oxford 1977.

D. Gabbay, *Semantical Investigations in Heyting's Intuitionistic Logic*, Reidel, Dordrecht 1981.

A.S. Troelstra, *Choice Sequences: A Chapter of Intuitionistic Mathematics*, Clarendon Press, Oxford 1977.

K. Segerberg, *An essay in Classical Modal Logic*, 3 volumi, Uppsala 1971.

J. van Benthem, *Modal Logic and Classical Logic*, Bibliopolis, Napoli 1983.

G.E. Hughes e M.J. Cresswell, *Introduzione alla logica modale*, a cura di Claudio Pizzi, Il Saggiatore, Milano 1983 [edizione originale 1973].

A.N. Prior, *Time and Modality*, Clarendon Press, Oxford 1968.

Id., *Past, Present and Future*, Clarendon Press, Oxford 1967.

AA.VV., *La logica del tempo*, a cura di Claudio Pizzi, Boringhieri, Torino 1974.

AA.VV., *Le logiche libere*, a cura di Ermanno Bencivenga, Boringhieri, Torino 1976.

E. Bencivenga, *Una logica dei termini singolari*, Boringhieri, Torino 1980.

H. von Wright, *An Essay in Modal Logic*, North Holland, Amsterdam 1951.

L. Åqvist, *Introduction to Deontic Logic and the Theory of Normative Systems*, Bibliopolis, Napoli 1987.

J.B. Rosser e A.R. Turquette, *Many-valued Logics*, North Holland, Amsterdam 1958.

AA.VV., *Selected Papers on Łukasiewicz sentential Calculi*, Ossolineum, Wroclaw 1977.

A.R. Anderson e N.D. Belnap jr, *Entailment*, volume I (1975), volume II (1992), Princeton University Press, Princeton.

P.J. Cohen, *La teoria degli insiemi e l'ipotesi del continuo*, con appendice di G. Lolli, Feltrinelli, Milano 1973 [edizione originale 1966].

G.H. Moore *The Origins of forcing*, in *Logic Colloquium '86*, North Holland, Amsterdam 1988, pp. 143-173.

J.L. Bell, *Boolean-valued Models and Independence Proofs in Set Theory*, Clarendon Press, Oxford 1977.

AA.VV., *Axiomatic Set Theory*, 2 volumi, a cura rispettivamente di D. Scott e T. Jech, AMS, Providence 1971, 1974.

P. Vopěnka e P. Hajek, *The Theory of Semisets*, North Holland, Amsterdam 1972.

P. Vopěnka, *Mathematics in the Alternative Set Theory*, Teubner, Lipsia 1979.

CAPITOLO SETTIMO
LE NUOVE PROSPETTIVE DEGLI ANNI SESSANTA

1. INSIEMI, DEFINIBILITÀ, COSTRUZIONI

1.1 *La teoria degli insiemi dopo Cohen*

Nel capitolo precedente abbiamo visto come l'introduzione del *forcing* e dei metodi ad esso collegati (modelli booleani, semisets, ecc.) abbiano completamente modificato il panorama delle indagini sulla teoria degli insiemi portando alla risoluzione di antichi problemi e permettendo di porne di nuovi e significativi. È proprio l'enorme ricchezza di questi risultati che rende impossibile anche il tentativo di offrire qui un panorama degli sviluppi degli studi insiemistici dopo i primi anni sessanta. Ci limiteremo quindi ad un rapidissimo cenno sulle direzioni principali di ricerca che si sono aperte dopo i lavori di Cohen rimandando a specifici testi (alcuni dei quali elencati in bibliografia) per una trattazione più adeguata.

Il punto di partenza è dato naturalmente dalle ricerche riguardanti l'assioma di scelta e l'ipotesi del continuo di cui appunto l'introduzione del *forcing* aveva permesso di dimostrare l'indecidibilità in $\mathfrak{ZF}$. Se fra i primi risultati ottenuti vanno citati i lavori del 1963 di Levy e Feferman sui rapporti tra varie forme di definibilità e l'assioma di scelta, decisiva fu la dimostrazione data da J. Halpern nel 1964 del fatto che il principio dell'ideale primo (*PIP*) non implica l'assioma di scelta ed è strettamente più debole pur se anch'esso non dimostrabile dagli assiomi di $\mathfrak{ZF}$. Il risultato rispondeva a una domanda che si era posta sin dal primo apparire del teorema e apriva la strada ad interessanti ricerche sulla possibilità di rimpiazzare con esso l'assioma di scelta in molte dimostrazioni in specifici settori della matematica.

Altro tema di interesse fondamentale riguarda il concetto stesso di finito. Ricerche sistematiche su questo argomento erano state condotte ad esempio da Tarski già nel 1924. Come si ricorderà, per Dedekind finito è ogni insieme che non sia equipotente ad una sua parte propria, mentre la definizione più consueta riconduce la finitezza di un insieme X al suo essere equipotente ad un segmento iniziale dei numeri naturali, al fatto cioè di avere un cardinale che è un numero naturale. Per provare l'equivalenza delle due nozioni (in particolare che se X è finito nel senso di Dedekind è finito anche nell'altro senso) era in generale usato *AS*; risultava così naturale chiedersi se principi più deboli fossero sufficienti allo scopo e per converso quali conseguenze di *AS* l'equivalenza permettesse di ottenere.

Che principi non troppo deboli siano necessari per provarla risultava dagli stessi contromodelli di *AS* costruiti da Mostowski per $\mathfrak{ZFA}$ in cui l'equivalenza è falsa: non stupisce quindi che forti si siano rivelate le assunzioni che permettono di ricavarla (come ad esempio il principio per cui non esiste catena discendente infinita di cardinali, introdotto da Hessenberg). Per converso, l'equivalenza ha conseguenze di estrema importanza, prima fra tutte – come provato da Kleinberg nel 1959 – il teorema di Ramsey (dimostrato originariamente usando *AS*). Si tratta – come si ricorderà – di una generalizzazione del principio dei cassetti di Dirichlet, di estrema importanza non solo nella combinatorica finita (come avremo occasione di vedere più avanti parlando dei modelli non standard dell'aritmetica) ma soprattutto in quella infinita che proprio dopo i primi anni sessanta ha avuto uno sviluppo impetuoso per opera in particolare di Paul Erdös e A. Hajnal.

Quanto detto sopra per *AS* e *PIP* si potrebbe ripetere per altri principi centrali nell'indagine non solo matematica, ma anche metamatematica, come il *principio PO dell'ordinamento* («ogni insieme ammette un ordinamento lineare»), l'assioma *ASD* delle *scelte dipendenti*, introdotto da P. Bernays nel 1949, secondo il quale se R è una relazione binaria su X tale che $\forall\, a \in X$ esiste un $b \in X$ per cui $R(a,b)$ allora esiste una funzione $f \subseteq R$, o l'assioma *ASN* delle *scelte numerabili*, che si ottiene da *AS* limitandosi a famiglie numerabili di insiemi. Nel 1969 Ronald B. Jensen dimostrava che *ASN* non implica *ASD*, mentre a Solovay si deve la dimostrazione che *ASD* è strettamente più debole di *AS* e a Mathias quella che *PIP*

non scende da *PO*. L'intergioco fra questi e altri principi è stato indagato in molteplici direzioni come pure la possibilità di ottenere teoremi metamatematici classici senza utilizzare principi come questi. In questo contesto ci limiteremo a ricordare la dimostrazione data da D. Pincus nel 1972 della indipendenza di *PIP* dal teorema di Hahn-Banach, teorema che – come mostrato nel 1955 da Łos e Ryll-Nardzewski – scende invece da *PIP* senza dover sfruttare tutta la forza di *AS*.

Questi risultati sono interessanti soprattutto in quanto nell'Analisi classica, ed in particolare nella teoria della misura, sono spesso *ASD* o *ASN* a intervenire piuttosto che il più forte *AS*. In questo contesto quest'ultimo assioma ha spesso conseguenze «negative», come ad esempio nel classico risultato di Vitali in base al quale esistono insiemi limitati di reali che *non* sono misurabili nel senso di Lebesgue. Uno dei risultati più interessanti della ricerca insiemistica degli anni sessanta riguarda proprio questi temi ed è la dimostrazione data da Solovay nel 1970 che, sotto l'ipotesi dell'esistenza di un cardinale inaccessibile, è coerente assumere accanto a $\mathfrak{ZF}$ e *ASD* che ogni sottoinsieme di R sia misurabile.

La domanda che si presenta naturale è allora se sia possibile trovare addirittura ipotesi *alternative* ad *AS* che ci permettano di ottenere i vantaggi di *ASD* o *ASN* nell'Analisi senza gli svantaggi di *AS*. Una proposta in questa direzione era stata fatta nel 1962 da Hugo Steinhaus e Jan Mycielski con *l'assioma di determinatezza AD*. Formulato in termini della teoria dei giochi l'assioma afferma che se *S* è un insieme di reali, cioè di successioni numerabili di 0 e 1 e se due giocatori scelgono alternativamente in $\{0,1\}$ formando così una successione infinita, il gioco sarà *determinato* nel senso che c'è una strategia che garantisce che la successione costruita è in *S* (oppure non è in *S*). Una delle ragioni che avevano condotto Mycielski ad accettare l'assioma (la cui *negazione* è conseguenza di *AS*) era che esso rendeva impossibili conseguenze paradossali di *AS* come la decomposizione di una sfera in due sfere ad essa congruenti provata, come si ricorderà, da Banach e Tarski. Ciò che Mycielski mostrava era che da *AD* non solo segue che ogni insieme di reali è misurabile, ma anche che *ASN* ristretto a sottoinsiemi di *R* è una conseguenza di *AD* (che ciò non valga per *ASD* è stato invece provato da Solovay nel 1978).

Se *AS* ha conseguenze sgradite, è innegabile però il suo ruolo in

molti settori della matematica astratta così che – almeno a partire da Cantor – è per molti difficile non vederlo come implicato dal concetto stesso di insieme *astratto*. Questo spiega allora come si siano cercati indebolimenti di *AD* che non fossero in contraddizione con *AS* pur mantenendone gli aspetti positivi. È questo il caso dell'assioma di determinatezza proiettiva *ADP* che si ottiene da *AD* restringendosi ad insiemi di reali che siano nella gerarchia proiettiva.

Gli insiemi *proiettivi* si ottengono da quelli *boreliani* (vale a dire quegli insiemi generati dagli aperti mediante complementi e unioni numerabili) applicando le due operazioni di complementazione e passaggio all'immagine diretta di funzioni continue. In questo senso essi costituiscono una collezione di sottoinsiemi di R semplici, in quanto definibili con mezzi limitati ed una misura di questa semplicità è data dal parallelismo con gli insiemi analitici della teoria della ricorsività di cui abbiamo parlato nel capitolo V. Come gli insiemi analitici, i proiettivi si dispongono in una gerarchia: Σ_1^1 (analitici) sono le immagini continue di boreliani, Π_n^1 i complementi di insiemi di Σ_n^1, infine Σ_{n+1}^1, le immagini continue di insiemi in Π_n^1.

La gerarchia riflette la complessità (in termini di quantificatori) della definizione degli insiemi in oggetto e non stupisce quindi che problemi di esistenza di strategie per giochi associati a sottoinsiemi di R appartenenti alla gerarchia ne rivelino proprietà significative. Le assunzioni di determinatezza comportano infatti (affermando la possibilità di strategie globali per *ogni* mossa dell'avversario, in presenza di strategie locali) l'invertibilità di quantificatori nei prefissi ed è questo collegamento che spiega l'importanza di considerazioni riguardanti giochi quando si abbia a che fare con prefissi ed in generale con problemi di definibilità. Collegamenti di questo tipo furono sistematicamente utilizzati all'interno della teoria dei modelli per linguaggi infinitari da Keisler nel 1969, ma già nel 1957 in Polonia – dove l'assioma di determinatezza era stato studiato per la prima volta – Ehrenfeucht ne aveva mostrato l'utilità formulando in questi termini le tecniche di Fraïssé degli isomorfismi locali.

Non è difficile provare che ogni gioco aperto è determinato e nel 1967 David Blackwell usò questo fatto per ridimostrare alcuni vecchi risultati di Lusin sugli insiemi analitici. Aperta rimaneva la

questione per giochi determinati da insiemi più complessi della gerarchia proiettiva. Si può provare che in $\mathfrak{ZF} + AS$ la determinatezza dei giochi analitici non è dimostrabile, mentre questo vale sotto l'ipotesi che esista almeno un cardinale misurabile, come mostrato da A. Martin nel 1970. Lo stesso Martin nel 1975 avrebbe ottenuto un altro risultato fondamentale al riguardo e cioè la determinatezza dei giochi boreliani, gli insiemi di complessità immediatamente minore di quelli analitici, senza assiomi ulteriori. Fu proprio questo fatto a rendere plausibile l'assunzione di forme intermedie di *AD* come *ADP* mediante la quale – come mostrato da Addison, Martin e Y. Moschovakis – era possibile ottenere una teoria pressoché completa della struttura degli insiemi proiettivi per *ogni* livello della gerarchia.

Tra le conseguenze più interessanti della determinatezza per ampie classi di insiemi di reali, c'è quella di godere della proprietà di Baire, di essere misurabili rispetto ad arbitrarie misure di Borel σ-finite ed infine la proprietà – se più che numerabili – di contenere sottoinsiemi *perfetti* (insiemi chiusi, non vuoti, che non contengono punti isolati, tali cioè che ogni loro intorno aperto abbia in comune con l'insieme solo il punto stesso). L'importanza di quest'ultima proprietà (nota appunto come *proprietà del sottoinsieme perfetto*) risulta evidente una volta che si ricordi il classico teorema di Cantor-Bendixon, in base al quale ogni insieme chiuso di reali si decompone univocamente in una somma disgiunta $P \cup S$ dove P è un insieme perfetto e S è numerabile. Ogni insieme perfetto – è facile dimostrare – ha cardinalità esattamente $2^{\aleph_0}$ così che corollario immediato del teorema è che per i chiusi vale l'alternativa posta dall'ipotesi del continuo: o sono numerabili o sono di cardinalità $2^{\aleph_0}$. Lo stesso vale per tutti gli insiemi di reali che godono della proprietà del sottoinsieme perfetto, così che possiamo concludere che *IC* vale per i boreliani e – nell'ipotesi di determinatezza proiettiva – vale per tutti i proiettivi. Il risultato si estende anche agli insiemi analitici (Σ^1_1-insiemi) di reali se assumiamo l'ipotesi *IM* dell'esistenza di un cardinale misurabile nella forma del risultato di Martin sopra citato. Il vantaggio di *ADP* su *IM* è che se il secondo principio ci permette di stabilire la validità di *IC* per gli analitici, non è in grado di dire nulla sui livelli alti della gerarchia dei proiettivi, cosa che invece fa *ADP*, permettendoci anche di stabilire *IC* per questi insiemi.

Siamo così giunti ad occuparci dello *status* dell'ipotesi del continuo e la cosa non sorprende se si pensa che sin dall'origine della teoria descrittiva degli insiemi il terreno naturale per saggiare la validità di *IC* è stato l'indagine degli insiemi definibili studiati dalla teoria descrittiva.

Se i risultati citati sopra sui boreliani (e che si possono estendere anche ai primi livelli della gerarchia proiettiva Σ^1_1 e Π^1_1 senza ulteriori ipotesi) sembrerebbero mostrare almeno la plausibilità di *IC* come generalizzazione di proprietà verificate per insiemi «semplici», a conclusioni diverse si arriva se si considera che *IC* si può formulare dicendo che ogni buon ordinamento di un insieme di reali avrà tipo d'ordine ω_2, il secondo cardinale più che numerabile. Si vede allora che la negazione di *IC* equivale ad affermare che esiste una suriezione (o funzione suriettiva)

$$f: \mathrm{R} \to \omega_2$$

dove ω_2 viene visto come l'insieme dei suoi predecessori. In generale ogni funzione

$$f: \mathrm{R} \to Ord$$

dà origine a un *pre-buon ordinamento* (prewellordering) di R considerando la controimmagine di f e bene ordinando le classi d'equivalenza rispetto alla relazione «avere la stessa immagine» in base agli ordinali associati; ora Martin e Solovay nel 1970 ottennero risultati che mostrano la coerenza e quindi anche la plausibilità del fatto che esistano pre-buoni ordinamenti semplici la cui lunghezza *contraddice* quella imposta da *IC* così che non sembrerebbe possibile ottenere indicazioni univoche sulla accettabilità di *IC* considerando insiemi di reali di struttura semplice.

La situazione – come negli anni immediatamente seguenti ai contributi di Cohen si realizzò ben presto – è ancora più drammatica per quanto riguarda l'aritmetica cardinale. Prima degli interventi di Gödel e di Cohen, l'unica indicazione nota sui rapporti tra $\aleph_\alpha$ e $2^{\aleph_\alpha}$ era il risultato (provato da J. König nel 1905) in base al quale $2^{\aleph_\alpha}$ non può avere *cofinalità* $\leq \aleph_\alpha$, vale a dire non può risultare somma di meno di $\aleph_\alpha$ cardinali tutti minori di $2^{\aleph_\alpha}$. Risultati congiunti di Cohen e Solovay mostrarono già nella prima metà degli anni sessanta che quella di König è essenzialmente l'*unica* re-

strizione che si può imporre al valore di $2^{\aleph_\alpha}$: per il resto tutto è possibile. La situazione è più complessa se consideriamo *IGC*. Sempre utilizzando la tecnica di contrazione dei cardinali, Easton mostrò nel 1970 che per $\alpha > 0$ anche $\aleph_\alpha$ e $2^{\aleph_\alpha}$ non sono vincolati a nessuna condizione se non quella di König tranne nel caso si abbia a che fare con cardinali *singolari*, cardinali κ cioè la cui cofinalità è minore di κ. Così ad esempio se esistevano modelli in cui $2^{\aleph_\alpha}$ poteva assumere valori qualsiasi compatibili con la condizione di König nel caso di $\aleph_\alpha$ *regolare* (cioè non singolare), non si riusciva a sapere se era coerente con $\mathfrak{ZF}$ $2^{\aleph_n} = \aleph_{n+1}$ o anche $2^{\aleph_n} > \aleph_{n+1}$.

Uno dei problemi generali della ricerca insiemistica dopo Cohen divenne così questo *problema dei cardinali singolari* cui lavorarono ricercatori come Karel Prikry, Jack Silver, Menachem Magidor ed altri. Solo verso il 1976 Silver risolse il problema per $\aleph_\alpha$ di cofinalità $> \aleph_0$ in una direzione che non si sarebbe aspettata, mostrando cioè che entro $\mathfrak{ZF}$ + *AS* senza ipotesi aggiuntive si può provare che se $\aleph_\alpha$ è singolare e di cofinalità $> \aleph_0$ e se $2^{\aleph_\beta} = \aleph_{\beta+1}$ per tutti i $\beta < \alpha$, allora $2^{\aleph_\alpha} = \aleph_{\alpha+1}$. Il risultato fu poi esteso da Fred Galvin e Hajnal, individuando specifici confini su $2^{\aleph_\alpha}$ per dati valori di $\aleph_\alpha$, ma nella sua generalità il problema è ancora irrisolto.

Prima di lasciare la questione dell'ipotesi generale del continuo è opportuno ricordare un altro classico problema riguardante la struttura dei reali che fu affrontato con successo in questo periodo generalizzando in modo essenziale le tecniche del *forcing*. Ci riferiamo al problema che Michael Suslin aveva posto nel 1920 e che riguarda la caratterizzabilità (a meno di isomorfismi) di R come un insieme ordinato $< R, \leqslant >$ senza estremi, completo (nel senso di Dedekind) e tale che ogni famiglia di intervalli propri a due a due disgiunti sia numerabile. L'*ipotesi di Suslin IS* afferma che R è così caratterizzabile, sostituendo così alla densità di Q in R la condizione (apparentemente più debole) sugli intervalli, detta anche delle anticatene numerabili.

Nel 1968 Ronald Jensen riuscì a costruire insiemi ordinati che soddisfano le condizioni di Suslin ma non risultano isomorfi a R assumendo l'assioma di costruibilità $V = L$, e nel 1967 Thomas Jech e Stanley Tennenbaum indipendentemente provarono la *non* dimostrabilità di *IS* anche in presenza di *IC*. Quello che però emerse ben presto era che anche il metodo del *forcing* nella versione originale di Cohen non era in grado di risolvere il problema

della indecidibilità o meno di *IS*. Riformulata l'ipotesi in termini di alberi (i cosiddetti *alberi di Suslin*), l'idea fu quella di considerare come insiemi di condizioni alberi come quelli di Suslin numerabili ed iterare la costruzione di estensioni generiche nel transfinito. Con questa tecnica Tennenbaum e Solovay riuscirono nel 1971 a provare che anche *IS* era coerente con *IC* e che quindi l'ipotesi di Suslin risultava indecidibile anche in presenza di quella del continuo. A parte l'interesse del problema in sé, la questione della coerenza dell'ipotesi di Suslin risulta importante anche perché è all'origine di altri sviluppi riguardanti l'estensione delle tecniche di Cohen, tecniche che portarono alla formulazione di un nuovo assioma, noto col nome di *assioma di Martin* (*AM*) che esplicitamente afferma l'esistenza di particolari filtri generici.

L'assioma ha conseguenze estremamente interessanti sul piano matematico in quanto – come mostrato in un classico lavoro dallo stesso Martin in collaborazione con Solovay – pur essendo coerente con la negazione di *IGC*, ha molte delle sue conseguenze e permette inoltre di risolvere problemi importanti: tra questi, non solo *IS* – di cui rivela in modo più diretto la non derivabilità dagli assiomi di $\mathfrak{ZF}$ – ma pure, come mostrato da Saharon Shelah nel 1974, il *problema di Whitehead* sui gruppi abeliani; utilizzando l'assioma di Martin, Shelah provò che esistono gruppi abeliani che pur soddisfacendo la condizione di Whitehead non sono liberi. Il risultato è tanto più interessante in quanto – come nel caso della negazione di *IS* – anche il problema di Whitehead ha soluzione *positiva* una volta che si assuma l'assioma di costruibilità $V = L$. Questo è solo un esempio di una lunga lista di ipotesi matematiche specifiche che non riguardano la nozione astratta di insieme, ma strutture ben determinate, di cui le tecniche iniziate da Cohen hanno permesso di mostrare l'indecidibilità in base agli assiomi ordinari. Soprattutto negli ultimi vent'anni risultati del genere si sono moltiplicati nella topologia generale, nella teoria dei gruppi e nell'Analisi funzionale.

Tutti questi fatti – oltre ovviamente al problema del continuo – portano naturalmente alla domanda di come sia possibile individuare *nuovi assiomi* con cui estendere l'ordinaria teoria degli insiemi. In astratto sono possibili due tipi di soluzione a questo interrogativo: la ricerca di assiomi che ci permettano di estendere la gerarchia di von Neumann ampliando l'universo, o al contrario

principi – come l'assioma di costruibilità – che se non l'altezza dell'universo (la lunghezza della scala degli ordinali) ne restringono almeno la larghezza, permettendoci di selezionare ad ogni passo solo alcuni dei sottoinsiemi del livello precedente. Non stupisce allora che una delle conseguenze dei risultati di Cohen sia stato il rinascere d'interesse per l'assioma di costruibilità ed in generale per l'universo dei costruibili.

Ricerche fondamentali a questo riguardo furono condotte soprattutto da Jensen che indagò sistematicamente la struttura fine dei costruibili portandone alla luce conseguenze dal punto di vista combinatorio. Sono state queste ricerche (sistematicamente sviluppate in seguito da altri ricercatori, tra cui Keith Devlin) a permettere quelle applicazioni dell'assioma a specifici problemi matematici di cui abbiamo parlato.

Se di grande interesse sono state le indagini – di cui abbiamo dato qualche indicazione nel capitolo precedente – sui rapporti tra costruibilità, teoria descrittiva e ricorsività generalizzata, il tema che forse ha avuto più sviluppi dopo le indagini di Cohen è stato però quello dei rapporti tra costruibilità e grandi cardinali.

L'assunzione di grandi cardinali e quindi di altezze sempre maggiori della gerarchia di von Neumann costituisce l'altra via naturale per individuare principi con cui rafforzare gli assiomi di $\mathfrak{ZF}$. L'idea che l'assunzione di assiomi dell'infinito sempre più forti (che garantiscono l'esistenza di cardinali sempre più grandi) possa permettere di decidere questioni che riguardano gli enti che si situano ai livelli bassi della gerarchia era stata avanzata da Gödel sostanzialmente già dagli anni trenta, in connessione con le sue indagini sulla lunghezza delle dimostrazioni, e riflette un atteggiamento ben preciso rispetto ai concetti insiemistici: i principi astratti non sono immediatamente evidenti ma possono essere giustificati in base al tipo di questioni che permettono di risolvere su oggetti più direttamente controllabili.

Che le due vie per «completare» gli assiomi insiemistici siano alternative risultava chiaro già nel 1961 quando D. Scott, utilizzando particolari ultraprodotti, mostrò come l'assioma di costruibilità fosse incompatibile con l'assunzione IM dell'esistenza di almeno un cardinale misurabile. Come già abbiamo avuto occasione di ricordare, un cardinale κ è misurabile quando esiste un insieme I di cardinalità κ per cui in $\mathscr{P}(I)$ si ha un ultrafiltro non principale

κ-completo, nel senso che ogni famiglia $\{X_i\}_{i\in J}$ di sottoinsiemi di I appartenente al filtro avrà intersezione ancora nel filtro se la cardinalità di J è minore di κ. Sappiamo dal capitolo precedente che i cardinali misurabili erano stati introdotti nel 1930 da Ulam in connessione col problema generale della misura, ed è facile provare che coincidono con quelli su cui è definita una misura σ-additiva non banale su **2** e che l'ultrafiltro della definizione è costituito dagli insiemi di misura 1. Ulam provava che i cardinali misurabili devono essere grandi almeno quanto il primo cardinale fortemente inaccessibile e lo stesso vale per i cardinali a misura reale se e solo se ogni cardinale debolmente inaccessibile è maggiore di $2^{\aleph_0}$.

I rapporti tra *IM* ed altri assiomi forti dell'infinito con l'ipotesi di costruibilità furono ampiamente indagati nei primi anni settanta da F. Rowbottom e soprattutto da Silver che nella sua tesi di dottorato del 1971 sviluppò quella tecnica degli *indiscernibili* in *L* che si fonda sulla generalizzazione di teoremi tipo quello di Ehrenfeucht e Mostowski sugli automorfismi di modelli, che si è rivelata di straordinaria utilità anche nello studio dei rapporti tra assiomi forti e teoria descrittiva degli insiemi. Questo tipo d'indagine era stato inaugurato da Solovay nel 1969 dimostrando che da *IM* scende che ogni insieme Σ^1_2 della gerarchia proiettiva gode della proprietà di Baire, è misurabile ed ha, se più che numerabile, un sottoinsieme perfetto. Questi come altri risultati, alcuni dei quali abbiamo già ricordato, portano in primo piano il problema dei rapporti tra gli assiomi forti e determinatezza, ed in particolare diventava fondamentale per tutta la teoria descrittiva il problema di sapere se *IM* implica la determinatezza per gli insiemi costruibili relativamente a R. Questa è la più forte ipotesi di determinatezza che possiamo fare per quanto riguarda la teoria descrittiva ed indagini importanti su questo tema sono state condotte in particolare da Martin.

Ad un minore ottimismo circa il ruolo degli assiomi forti come principi in grado di decidere questioni lasciate indecise da $\mathfrak{Z}\mathfrak{F}$ inducono invece alcuni risultati di Levy e Solovay che risalgono al 1968 e mostrano come gli assiomi forti non pongano ostacoli alla costruzione di estensioni generiche. Era questo un sospetto già espresso da Cohen in contrasto con le speranze di Gödel e che ha portato a interessanti ricerche sulle proprietà di grandi cardinali preservate dalle estensioni generiche e quindi non sottoponibili a

collasso. Così, ad esempio, Solovay ed indipendentemente Mc Aloon hanno dimostrato che i cardinali misurabili vengono conservati da particolari estensioni generiche e lo stesso vale per quelli compatti.

Ciò non toglie ovviamente che implicita nell'idea stessa di insieme astratto sia la possibilità di iterare le costruzioni della gerarchia di von Neumann e superare ogni limitazione considerando come insiemi gli universi definiti per poi procedere ad ulteriori costruzioni. Lo studio degli assiomi forti dell'infinito è stato quindi sin dal tempo di Zermelo un campo attivo di ricerca che ha portato ad isolare – come avremo occasione di osservare parlando di linguaggi infinitari – cardinali debolmente e fortemente inaccessibili, cardinali di Mahlo, iperinaccessibili, ecc. Ruolo importante in questi processi di generazione di insiemi sempre più grandi hanno svolto considerazioni riguardanti i *punti fissi* delle specifiche costruzioni come pure i *principi* di riflessione che – partendo dalle considerazioni che abbiamo ricordato a suo tempo parlando dei *teoremi* di riflessione – cercano di individuare cardinali sempre più grandi sfruttando l'idea che l'assunzione che date proprietà dell'universo si riflettano in un livello della gerarchia di von Neumann implica l'esistenza nell'universo stesso di ordinali «grandi» che danno un confine all'altezza del livello in esame.

Ricerche interessanti a questo riguardo furono condotte in tempi diversi da Bernays, Levy, Vaught e Montague e tanti altri e non deve stupire che esse si siano intensificate dopo i risultati di indipendenza di Cohen. Un aspetto importante degli assiomi forti così individuati è infatti che, per usare l'espressione di Solovay, «la loro verità implica la loro indipendenza» nel senso che sulla base della loro esistenza risulta dimostrabile la coerenza del sistema d'assiomi cui sono stati aggiunti e, quindi, per il teorema di Gödel, non risultano essi stessi dimostrabili, sono cioè autentici assiomi *aggiuntivi* e permettono di provare nuovi principi aritmetici. Il tentativo di sfruttare al massimo l'idea che sta al fondo dei principi di riflessione ha portato ad indagini estremamente interessanti che coinvolgono l'esistenza di immersioni elementari tra universi, la stessa distinzione tra classi e insiemi e nozioni modali.

Lavori su questi temi sono stati condotti nei primi anni settanta da W. Powell e William Reinhardt che hanno mostrato come l'idea di riflessione comporti in sé la attualizzazione di esistenze pos-

sibili e quindi il ricorso a distinzioni di tipo modale. In questo contesto non stupisce la ripresa di interesse per altri tipi di assiomatizzazione della teoria degli insiemi come quella presentata da W. Ackermann nel 1965, che si basa appunto su una nuova analisi della distinzione cantoriana tra coerente e non. L'idea di Ackermann è che perché una collezione di insiemi sia un insieme è necessario e sufficiente che «quello che appartiene e quello che non appartiene alla collezione sia delimitato in modo netto». Ackermann assumeva così nel linguaggio, accanto a $\in$, una costante predicativa M («essere un insieme») ed oltre all'assioma di estensionalità e di costruzione di classi postula uno schema (lo *schema di Ackermann*) che sostanzialmente afferma che se tutti gli oggetti che godono della proprietà A (che non deve però coinvolgere la costante M) sono insiemi, allora anche l'estensione della proprietà sarà un insieme. A questi assiomi ne andava aggiunto un altro che stabilisce che ogni elemento ed ogni sottocollezione di un insieme è un insieme. Il sistema di Ackermann si presentava come particolarmente potente e come tale fu studiato sin dal 1969 da Levy e da R. Grewe. Il fatto nuovo venne nel 1970 quando Reinhardt dimostrò che – fatte alcune precisazioni – esso non è più forte di $\mathfrak{ZF}$, ma gli è sostanzialmente equivalente.

In questa stessa prospettiva di ripensamento delle assunzioni di fondo va visto anche il rinascere di interesse per sistemi che partono da analisi diverse da quella zermeliana di insieme, come quelli presentati da Quine a partire dagli anni quaranta. I sistemi di Quine $\mathfrak{ML}$ (*Mathematical Logic*) e $\mathfrak{NF}$ (*New Foundations*) si fondano su un ripensamento della gerarchizzazione in tipi proposta da Russell e ne forniscono una versione per così dire relativizzata mediante il concetto di *stratificazione*. Grosso modo l'idea di Quine è che si può evitare una gerarchizzazione *a priori* e per restringere l'assioma di comprensione è sufficiente postulare che la formula la cui estensione vogliamo assumere come insieme sia stratificata, ammetta cioè tra le sue variabili una gerarchizzazione in livelli per cui non si abbia mai che, se $x \in y$, x abbia livello maggiore o uguale a quello di y. La possibilità di diverse stratificazioni al posto di un'unica assegnazione di tipo, fa sì che i sistemi di Quine siano difficilmente confrontabili con $\mathfrak{ZF}$ o $\mathfrak{NBG}$, per cui vale un teorema di gerarchizzazione come quello di von Neumann.

Prova della radicale diversità dei sistemi di Quine da quelli

ordinari è ad esempio il fatto che a tutt'oggi il problema centrale di $\mathfrak{NF}$ è quello di provarne la coerenza; a ciò si aggiunga il risultato di Specker del 1953 sulla contraddittorietà di $\mathfrak{NF}$ con *AS*: questo contrasta con il risultato di Gödel che, come si ricorderà, stabilisce la coerenza di $\mathfrak{ZF}$ con l'assioma in questione, anche se, come ha dimostrato Jensen nel 1968, è possibile modificare $\mathfrak{NF}$ in modo da renderlo coerente con *AS*. Al di là di tutte le difficoltà, l'interesse fondamentale di $\mathfrak{NF}$ sta comunque nella sistematica adozione di quella *ambiguità dei tipi* che già i *Principia* avevano utilizzato. Come provato da Specker, infatti, $\mathfrak{NF}$ è coerente se e solo se lo è la teoria semplice dei tipi con l'aggiunta di assiomi che permettono "innalzamenti" ed "abbassamenti" uniformi di tipo nelle formule.

I principi di riflessione non sono però l'unico tipo di considerazione che porta ad individuare grandi cardinali. Un esempio di come altri tipi di idee possano fare questo è fornito dai cardinali misurabili e da quelli compatti e debolmente compatti di cui abbiamo già parlato. In questo caso a scattare è l'analogia e l'idea di generalizzare a cardinali più che numerabili proprietà di cui gode oppure non gode ω ma la famiglia dei cardinali accessibili (i singolari oppure quegli α per cui esiste β tale che $\beta < \alpha \leq 2^{\beta}$). Se inaccessibili, misurabili e fortemente compatti si presentano generalizzando proprietà di ω, i cardinali che si individuano nel secondo caso sono stati studiati a fondo da Tarski in collaborazione con diversi ricercatori, in particolare Erdös e Keisler (quest'ultimo pubblicò con Tarski, nel 1964, il fondamentale *From accessible to inaccessible numbers* (*Dai numeri accessibili agli inaccessibili*)).

Queste e altre considerazioni hanno portato a un panorama estremamente complesso per quanto riguarda l'esistenza di grandi cardinali, i loro rapporti di grandezza, le loro proprietà, ecc. In questo contesto, come abbiamo avuto occasione di ricordare, un grosso ruolo ha il concetto di immersione elementare tra universi parziali e la possibilità quindi di simulare un livello della gerarchia con altri. L'utilità di una simile impostazione è stata ampiamente dimostrata nel 1978 da Solovay in collaborazione con Reinhardt e Kanamori e non stupisce d'altra parte che in questo clima Scott nel 1974 abbia mostrato come possa essere fecondo parlare esplicitamente di livelli già nella descrizione degli universi d'insiemi ed assiomatizzarne direttamente le proprietà.

Questo progressivo complicarsi del quadro di riferimento con tutti i problemi e le alternative che lascia aperti, se da una parte mostrava la vitalità della ricerca insiemistica dopo l'intervento di Cohen, non poteva dall'altra non porre grossi interrogativi per quanto riguarda la sua rilevanza nei confronti della pratica matematica nel suo complesso. Il presentarsi di alternative radicali, la difficoltà di individuare metodi per scegliere tra ipotesi alternative, pone in dubbio l'idea stessa della teoria degli insiemi come base e fondamento di tutta la matematica. A questo si aggiunga il fatto che in non pochi casi gli sviluppi decisivi della matematica di quest'ultimo mezzo secolo (primo fra tutti il diffondersi di metodi della teoria delle categorie nella geometria, nell'algebra, nella topologia, ecc.) mostravano come sul terreno del concreto operare matematico altre concettualizzazioni, diverse da quella insiemistica, si mostrassero più efficaci, concettualizzazioni che a loro volta solo difficilmente e con innaturali ricostruzioni potevano essere riassorbite entro orizzonti insiemistici.

Interessanti a questo riguardo sono i vari tentativi di portare entro una fondazione insiemistica i metodi della teoria delle categorie che si succedettero a partire dalla fine degli anni sessanta. Implicito nella pratica di chi si serve delle categorie è il passare dalle categorie stesse (che per lo più sono classi proprie) a funtori che agiscono su categorie, a trasformazioni naturali tra funtori, ecc. Come è possibile parlare, nell'universo degli insiemi, di classi, di operazioni su classi, di operazioni tra operazioni su classi? Accanto a diversi – e sfortunati – tentativi di sviluppare una teoria delle classi di classi, si provarono diverse vie che andavano dall'uso di principi di riflessione alla relativizzazione ad opportuni universi (gli universi proposti da Grothendieck a questo scopo) ecc. Malgrado in questi tentativi si siano cimentati logici come Feferman, Kreisel, Engeler e, in Italia, Gabriele Lolli, non si può dire che gli esiti siano stati felici non tanto sul piano formale, quanto su quello concettuale.

L'approccio categoriale è decisamente eterogeneo a quello insiemistico e non stupisce allora che dopo diversi tentativi di Mac Lane, nel 1964 William Lawvere abbia tentato una fondazione della teoria delle categorie su se stessa presentando la categoria delle categorie come fondazione dell'intera matematica. Se questi tentativi si basavano su una visione ancora modellata sull'idea

classica di fondazione, la situazione è cambiata agli inizi degli anni settanta con la creazione della teoria elementare dei topos per opera di Lawvere e di Tierney. Non solo il concetto di topos risulta autonomo rispetto a quello d'insieme, ma – oltre a permettere un approccio più diretto allo studio dei metodi categoriali nel resto della matematica – è in grado di generalizzare la nozione stessa di universo degli insiemi, che risulta un particolare topos. Il fatto non è estrinseco e senza conseguenze in quanto, come sempre più si è andato vedendo, l'uso del *forcing*, i modelli booleani e molte loro generalizzazioni risultano casi particolari di tecniche che possono essere affrontate nella loro generalità proprio con metodi categoriali. Al di là quindi dell'opzione di base – insiemi o categorie? – rimane il fatto che, a partire dai lavori di A. Joyal, W. Mitchell, D. Higgs ed altri, dai primi anni settanta uno dei metodi più fecondi per studiare la teoria degli insiemi è divenuto proprio l'uso delle categorie.

Se le ricerche di tipo categoriale pongono in luce la problematicità dei rapporti tra teoria degli insiemi e nuovi sviluppi della matematica, ad analoghi ripensamenti la situazione creatasi dopo i risultati di Cohen ha portato rispetto alla matematica tradizionale, quella che non assume direttamente come oggetti di studio insiemi e strutture astratte, ma fa ricorso a queste solo nella misura in cui esse permettono di meglio studiare numeri naturali, reali, complessi, ecc. La domanda di fondo in questa prospettiva diviene quella di sapere entro che limiti è necessario postulare l'armamentario concettuale della teoria astratta degli insiemi e quanto invece non sia possibile evitare gli scogli della indecidibilità restringendo il proprio universo, più che cercando di estenderlo. È questa la seconda via di cui abbiamo parlato sopra e che ha trovato il suo rappresentante più significativo in Harvey Friedman.

Dietro la posizione di Friedman non sta alcun tipo di atteggiamento costruttivo, ma semplicemente la constatazione che la maggior parte degli oggetti della matematica classica si trovano a livelli bassi della gerarchia di von Neumann, hanno rango $< \omega + \omega$ e quindi diviene essenziale – se *questo* è il nostro oggetto d'indagine – domandarsi entro che limiti e dove risulta essenziale salire a livelli più alti, muoversi in quella che Friedman chiama la teoria degli insiemi superiore. Si tratta di una messa in discussione (non, si badi bene, di un rifiuto) degli sviluppi astratti della teoria degli insiemi,

che non a caso assume come punto di partenza l'originario sistema di Zermelo, $\mathfrak{Z}$, che coincide essenzialmente con $\mathfrak{ZF}$ meno l'assioma di rimpiazzamento: per avere gli oggetti di rango $< \omega + \omega$ non è infatti necessario postulare questo assioma.

La domanda è se esso o ulteriori principi più forti non si rendano necessari per studiare le proprietà degli oggetti a questi livelli. Non stupisce che in quest'ottica divengano centrali i concetti della teoria descrittiva i cui oggetti coincidono appunto con quegli enti «semplici» di cui si occupa la matematica classica. In un suo lavoro del 1971 intitolato appunto *Higher Set Theory and Mathematical Practice* (*Teoria superiore degli insiemi e pratica matematica*) Friedman sottolineava come la difficoltà nel trovare enunciati su oggetti matematici «semplici» (di rango $< \omega + \omega$) indipendenti da $\mathfrak{Z}$ ma dimostrabili in $\mathfrak{ZF}$ o rafforzamenti, stava nel fatto che questi enunciati, come spesso l'esperienza mostrava, implicavano in $\mathfrak{Z}$ l'esistenza di insiemi non costruibili e questo significava dover affrontare il problema – delle cui difficoltà abbiamo dato prove più sopra – di sapere quali principi accettabili esistono che contraddicono l'ipotesi di costruibilità.

Friedman considerava come esempio di enunciato matematicamente significativo la determinatezza degli insiemi boreliani di cui mostrava la non dimostrabilità in $\mathfrak{Z}$ e poneva il problema di sapere se esso risultava dimostrabile in $\mathfrak{ZF}$. Come sappiamo, questo è vero – come fu provato da Martin nel 1975 – così che l'enunciato in questione costituisce probabilmente il primo esempio di teorema che non comporta banalmente l'assioma di rimpiazzamento (e quindi la teoria superiore degli insiemi) che può essere dimostrato solo mediante tale assioma. A partire dagli anni settanta Friedman avrebbe sviluppato il suo programma volto a individuare i contesti della matematica classica in cui principi astratti si rivelano necessari in svariate direzioni che l'avrebbero portato a introdurre tecniche per provare l'indipendenza diverse da quelle di Cohen che si applicano alla teoria superiore e la cui applicabilità, come sappiamo, cessa una volta che, conforme all'atteggiamento di fondo che abbiamo illustrato, non si esclude a priori la validità dell'assioma di costruibilità.

Avremo occasione di tornare più avanti su alcuni sviluppi di queste indagini parlando della *Reverse mathematics*; quel che vogliamo sottolineare ora è l'interesse che a partire dagli anni sessanta si

è andato sempre più accentuando per sistemi più deboli di $\mathfrak{ZF}$ e per tipi di universi che, come i costruibili, siano vincolati da considerazioni di definibilità. Già nella sua famosa conferenza di Princeton del 1946, Gödel aveva mostrato come i costruibili non fossero l'unico tipo di universo cui si potesse giungere partendo da condizioni di definibilità e come fosse plausibile considerare l'universo *OD* degli insiemi definibili in termini di ordinali, quegli insiemi cioè che possono essere definiti da espressioni che oltre alle costanti logiche e la quantificazione su insiemi possono contenere nomi per ordinali.

La differenza con i costruibili sta essenzialmente nel fatto che nel caso di *L* non possiamo quantificare su insiemi in generale, così che *OD* include *L* ma non coincide con esso. Fatto importante, anche la definibilità in termini di ordinali è una nozione assoluta ed è questo uno dei principali motivi del suo interesse. Come congetturato da Gödel, la sottoclasse *HOD* degli insiemi ereditariamente definibili in termini di ordinali (gli elementi di *OD* i cui elementi siano ancora nella classe) è modello non solo di $\mathfrak{ZF}$, ma anche di *AS*, come fu provato da J. Myhill e Scott nel 1971. Dalla loro introduzione *OD* e *HOD* sono stati fatti oggetto di diverse ricerche che ne hanno posto in luce similarità e differenze con *L*. Nel 1966 ad esempio Kenneth Mac Aloon ha provato che l'ipotesi che $V = HOD$ è coerente tanto con la negazione di *IC* quanto con l'assioma *IM* sull'esistenza di almeno un cardinale misurabile.

L'interesse per problemi di definibilità porta in primo piano i legami con la teoria generalizzata della ricorsività e con quella dei linguaggi infinitari su cui torneremo più avanti, e in questo contesto i problemi di assolutezza hanno un grande significato. L'elemento nuovo che emerge in modo sempre più netto negli anni sessanta è dato dall'analisi del significato dell'assolutezza in termini della forma delle formule che ci permettono di definire la gerarchia costruibile. Si può dimostrare infatti che l'assolutezza della nozione in esame dipende dal fatto che le formule di cui ci serviamo per definire la gerarchia sono del tipo $\Delta_1^{\mathfrak{ZF}}$ (se ci muoviamo in $\mathfrak{ZF}$) sono cioè formule che rispetto a $\mathfrak{ZF}$ sono equivalenti a formule tanto del tipo $\forall x_1,...,x_n \mathscr{A}$ quanto del tipo $\exists x_1,...x_n \mathscr{B}$, dove $\mathscr{A}$ e $\mathscr{B}$ sono formule senza quantificatori illimitati: appunto per questa persistenza rispetto alle estensioni e alle restrizioni le formule Δ_1 sono assolute.

L'idea di considerare la gerarchia delle formule insiemistiche in base ai blocchi di quantificatori illimitati è, come si ricorderà, di Azriel Levy (1965) e permette di vedere molto bene i rapporti tra costruibilità, proprietà delle gerarchie e generalizzazioni della ricorsività. Come si è visto infatti le gerarchie della teoria della ricorsività si basano in larga misura sulla complessità dei prefissi ed è questo che consente di estendere molti concetti della teoria della ricorsività basata sulla teoria dei numeri naturali a una teoria della ricorsività basata sulla teoria degli insiemi. Centrale è il problema di individuare il nucleo minimale della teoria degli insiemi che sia in grado di permetterci la definizione degli analoghi insiemistici delle funzioni ricorsive primitive e di dimostrare la chiusura rispetto a schemi definitori quali quello di recursione.

Il nucleo fondamentale individuato è costituito dalla teoria $\mathfrak{KP}$ di Kripke-Platek che si ottiene da $\mathfrak{ZF}$ eliminando gli assiomi della potenza e dell'infinito e restringendo lo schema di rimpiazzamento a formule Δ_0 (cioè senza quantificatori illimitati). $\mathfrak{KP}$ è sostanzialmente una versione in termini di appartenenza della teoria di Peano dei numeri naturali ed il ruolo dell'assioma di induzione è svolto dallo *schema* di fondazione (un principio di fondazione relativo a formule e non a insiemi). Entro $\mathfrak{KP}$ si possono così definire funzioni primitive ricorsive su insiemi e studiare le gerarchie che ne risultano. Fatto fondamentale, $\mathfrak{KP}$ si presenta come adeguato sfondo per lo studio della metateoria sintattica e anche semantica della stessa teoria degli insiemi. In particolare $\mathfrak{KP}$ ci permette di isolare una classe di insiemi notevolmente interessanti, quella degli insiemi *ammissibili*, i modelli appunto di $\mathfrak{KP}$, che costituiscono nelle generalizzazioni della teoria della ricorsività gli analoghi dell'insieme dei numeri naturali.[1]

Un aspetto particolarmente importante di questo tipo di ricerche, come vedremo brevemente più avanti, sta nella luce che esse gettano sulle proprietà semantiche e sintattiche dei linguaggi infi-

[1] Come si ricorderà, Kripke aveva isolato la nozione di *ordinale* ammissibile, partendo da una generalizzazione del calcolo di Gödel-Kleene delle funzioni ricorsive (cui si accennerà più avanti). L'idea di considerare gli *insiemi* ammissibili – che sono generalizzazioni degli insiemi L_α, con α *ordinale* ammissibile – e l'osservazione della validità per essi di uno schema particolare di riflessione è di Richard Platek, cui si deve propriamente la teoria della ricorsività su *insiemi* ammissibili.

nitari: come la teoria delle funzioni ricorsive numeriche è la base per la sintassi dei linguaggi finitari, così la teoria delle funzioni ricorsive su insiemi ammissibili è la base della sintassi dei linguaggi infinitari e il rilievo che in questo contesto acquistano gli ammissibili sta nel fatto che – come ricordato nel capitolo VI parlando del teorema di Barwise – per essi vale una forma del teorema di compattezza.

Lavori molto interessanti sugli insiemi ammissibili si devono a Barwise, Shoenfield, Gandy e altri che ne hanno indagato sistematicamente le proprietà, non solo in collegamento con i linguaggi infinitari, ponendo in luce un possibile modo di generalizzare i concetti di finito, ricorsivo, ecc. Nella misura in cui i concetti fondamentali della teoria ordinaria della ricorsività (ORT) si possono giustificare in base a considerazioni diverse (macchine di Turing, sistemi equazionali, definibilità all'interno di teorie, ecc.) è chiaro d'altra parte che saranno possibili diverse vie per raggiungere generalizzazioni adeguate. È di questo che ci occuperemo nel prossimo paragrafo.

1.2 *La generalizzazione del concetto di effettivo*

I primi tentativi di generalizzazione della ORT risalgono essenzialmente allo studio dell'analogia esistente fra le gerarchie aritmetica e analitica inaugurato da Kleene nel 1955. L'analogia stava nel fatto che si può considerare l'insieme dei predicati iperaritmetici come la generalizzazione al secondo ordine dell'insieme dei predicati ricorsivi (anche se essa non è così stretta come può sembrare: si veda più oltre il discorso sugli insiemi «metafiniti»). Si studiavano così i rapporti fra le gerarchie basate su tali insiemi di partenza ma ci si rese conto ben presto che l'analogia non era letterale e che all'insieme dei predicati ricorsivamente enumerabili (Σ_1^0) doveva corrispondere l'insieme Π_1^1, e *non* l'insieme Σ_1^1. L'estensione del concetto di funzione parziale ricorsiva poteva essere agevolmente ottenuta per esempio introducendo un opportuno sistema di indici per gli insiemi Π_1^1 e definendo le funzioni parziali Π_1^1 a partire da *grafi* Π_1^1: in questo modo alle funzioni ricorsive totali corrispondevano le funzioni iperaritmetiche, cioè quelle funzioni il cui grafo è Δ_1^1. Lo studio di questo tipo di analogia portò a

una serie di risultati assai interessanti non solo per l'ampliamento di conoscenza sulla gerarchia analitica, ma anche sulla funzione, spesso ambigua, dei concetti basilari della ORT.[2]

In relazione a questa tematica e allo studio dei funzionali di tipo finito, nel 1962 venne presentata da Kleene una generalizzazione delle macchine di Turing basata sull'«halting problem». Accanto agli usuali passi della procedura Kleene supponeva la possibilità di un passo «speciale» che consisteva nello stabilire le condizioni di stop di una macchina di Turing di cui è dato l'indice: tale passo è lecito però solo nel caso in cui la macchina di cui si studia la procedura o è ordinaria o contiene passi speciali di complessità inferiore (si tratta quindi di una definizione per induzione). Le funzioni computate da queste macchine risultavano essere esattamente le Π^1_1: quindi era possibile caratterizzare queste funzioni come quelle per il cui calcolo si può procedere in modo deterministico sulla base di un insieme finito di istruzioni, purché si supponga la capacità, ad ogni istante, di «passare in rassegna» una successione infinita di passi. Risulta così possibile descrivere tali macchine mediante un *albero* in cui al primo livello compaiono i passi della procedura, e ad ogni livello successivo i passi delle sottoprocedure che corrispondono ai passi speciali del livello precedente: ad ogni albero di questo genere veniva associato in modo naturale un ordinale che ne misura la complessità e che risultava essere un ordinale ricorsivo.

A questo punto era abbastanza naturale concepire una nuova generalizzazione della ORT, in cui l'insieme N dei numeri naturali era sostituito dall'insieme *Or* degli ordinali ricorsivi (più preci-

[2] Per chiarire meglio questo fatto possiamo per esempio notare che nell'estensione di un asserto (teorema) dalla gerarchia aritmetica a quella analitica, possiamo studiare l'analogia in senso stretto, trasformando cioè *ogni* concetto aritmetico nel suo analogo analitico, oppure operare solo parzialmente tale trasformazione (si ricordi che la gerarchia aritmetica è una sottogerarchia di quella analitica). Se ad esempio vogliamo definire il concetto di insieme Π^1_1-produttivo, possiamo stabilire che la funzione parziale produttiva ad esso associata sia Π^1_1-parziale (analogia totale) oppure che essa rimanga ancora parziale ricorsiva (analogia parziale) e i risultati che si potranno ottenere saranno ovviamente diversi: se l'estensione di un teorema vale nel caso dell'analogia totale si è portati a ritenere che la funzione dei concetti che compaiono nel suo asserto (nella ORT) è, per così dire, relativa all'universo in cui ci si muove; viceversa nel caso dell'analogia parziale si metterebbe in luce il valore assoluto (che non cambia estendendo l'universo) di alcuni di questi concetti.

samente *Or* è un particolare insieme di *notazioni* per gli ordinali ricorsivi, quindi ancora un insieme di numeri naturali): va notato che *Or* è un insieme Π^1_1 ed è completo nel senso che ogni insieme Π^1_1 è «riducibile» (sulla base di un'opportuna estensione del concetto di riducibilità della ORT) a *Or*. Questa generalizzazione, detta *metaricorsività*, fu proposta da Kreisel e Gerald Sacks nel 1965, sullo spunto degli studi condotti da Takeuti negli anni cinquanta. Un insieme (di ordinali ricorsivi) verrà detto: *metaricorsivamente enumerabile* se l'insieme di notazioni dei suoi elementi è Π^1_1; *metaricorsivo* se sia l'insieme delle notazioni dei suoi elementi sia il suo complemento (rispetto a *Or*) è Π^1_1; *metafinito* se l'insieme delle notazioni dei suoi elementi è Δ^1_1. Particolarmente interessante la generalizzazione del concetto di *finito*: esso viene interpretato come «ricorsivo e limitato» e nella generalizzazione ciò diventa «metaricorsivo e limitato (da un ordinale ricorsivo)», il che è equivalente a richiedere che sia Δ^1_1. Ciò, in connessione con il fatto che ovviamente la gerarchia analitica può essere studiata come una particolare restrizione della metaricorsività (identificando i numeri naturali con gli ordinali finiti, gli insiemi Π^1_1 coincidono con i metaricorsivamente enumerabili), indica come nell'analogia fra gerarchia aritmetica e gerarchia analitica non sempre sia opportuno far corrispondere gli insiemi Δ^1_1 agli insiemi ricorsivi, ma agli insiemi finiti.

Vi è inoltre una conseguenza assai importante: mentre ogni intersezione di un insieme numerico con un qualunque insieme numerico finito è ovviamente finita, non è detto che ogni intersezione di un insieme *A* di ordinali ricorsivi con insiemi metafiniti sia a sua volta metafinita. Se ciò si verifica in ogni caso, diremo che l'insieme *A* è *regolare*. Molto spesso si verifica che le estensioni dei teoremi sugli insiemi ricorsivamente numerabili agli insiemi metaricorsivamente enumerabili valgono solo supponendo tali insiemi regolari: d'altra parte si dimostra che ogni insieme metaricorsivamente enumerabile è «riducibile» a qualche insieme metaricorsivamente enumerabile e regolare. Si vede così come una delle conseguenze dei lavori di generalizzazione fu quella di mettere in luce più profondamente i nessi che intercorrono fra i vari concetti enucleati nell'ambito non generalizzato, nessi talora ambigui e oscuri proprio a causa della relativa «semplicità» o anche «concretezza» della struttura di partenza.

A livello di esposizione abbiamo presentato la metaricorsività come un'ulteriore generalizzazione dello studio della gerarchia analitica: anche se i rapporti sono molto stretti e lo studio della gerarchia analitica con i metodi della metaricorsività si rivela assai fruttuoso, tuttavia storicamente le basi della metaricorsività si svilupparono nell'ambito delle ricerche sul concetto di definibilità (al secondo ordine). Il collegamento può essere reso intuitivo considerando come alla «complessità» relativa di ogni particolare insieme che si vuol definire corrisponde la complessità di una formula in un opportuno linguaggio (in generale infinitario); questa a sua volta può essere analizzata mediante un opportuno albero (che potrebbe venir fatto corrispondere a una macchina di Turing generalizzata) e quindi un opportuno numero ordinale (o meglio una notazione per tale numero). Da questo punto di vista assumeva allora ovvia importanza lo studio degli insiemi numerici definiti in modo induttivo (Clifford Spector, 1961) in un linguaggio del primo ordine (sostituendo così a una definizione *esplicita* al secondo ordine una definizione *induttiva*, e quindi implicita, al primo ordine) e gli insiemi induttivi risultavano essere esattamente i Π^1_1.

Possiamo dire che in larga misura le ricerche condotte a partire dagli anni sessanta si presentano come generalizzazioni dei filoni precedentemente indicati. In particolare, quindi, schematizzando:

– lo studio approfondito delle generalizzazioni delle analisi basate sul concetto di «riducibilità» (gradi, ecc.) relativamente ad enti di tipo superiore a quello dei numeri naturali (funzionali oppure ordinali ricorsivi); un'ulteriore generalizzazione si basa sul concetto di *ordinale ammissibile* (intuitivamente un ordinale α è ammissibile se ogni «deduzione» a partire da un sistema finito di equazioni e che contenga cifre che siano notazioni di ordinali minori di α ha una «complessità» minore di α) e viene sviluppata in modo del tutto analogo (per quello che riguarda la struttura di base) alla metaricorsività: l'importanza di tale generalizzazione sta nel fatto che *comunque* sia dato un ordinale ammissibile, l'insieme degli ordinali minori si comporta (dal punto di vista della generalizzazione) come l'insieme dei numeri naturali (si noti che ogni cardinale è ammissibile):

– una parallela ripresa delle ricerche sulle definizioni induttive, non più solo in relazione alla struttura dei numeri naturali, ma generalizzate a strutture arbitrarie. A questo proposito ricordiamo

come questi studi si siano inizialmente sviluppati in un ambito alquanto diverso, che tendeva a generalizzare a strutture arbitrarie proprio il momento «calcolistico» della ORT: anche se i risultati ottenuti erano promettenti, l'apparato definitorio risultava complesso e spesso artificioso; il cambiamento di prospettiva, con il quale si è focalizzato l'interesse sulle *definizioni induttive*, ha permesso di inquadrare i risultati precedentemente raggiunti in una sistemazione assai più semplice dal punto di vista concettuale. Le ricerche in questo campo hanno permesso di ottenere nuovi risultati nell'ambito della generalizzazione basata sugli ordinali ammissibili cui si è fatto cenno al punto precedente. Lavori importanti in questa direzione sono dovuti a Y. Moschovakis che nel 1964 ha pubblicato un volume, *Elementary Induction on Abstract Structures* (*Induzione elementare su strutture astratte*) in cui si mostrano i collegamenti fra la teoria delle definizioni induttive, quella degli insiemi ammissibili e in generale lo sfondo insiemistico e di teoria dei modelli di questo approccio alla ricorsività. Collegate alle ricerche di Moschovakis sono quelle di Barwise, che nel suo volume *Admissible Sets and Structures* (*Insiemi ammissibili e strutture*) del 1975 ha sistematicamente posto in luce la centralità del concetto di definibilità per quanto riguarda le definizioni induttive e la ricorsività generalizzata, lo studio degli insiemi ammissibili e più in generale dei modelli della teoria degli insiemi, come pure la teoria dei modelli dei linguaggi infinitari.

Un altro campo di ricerca riguarda le sistemazioni di *tipo assiomatico* (come tali quindi applicabili a strutture entro certi limiti arbitrarie) che hanno il loro punto di partenza nella teoria degli indici della ricorsività ordinaria (teoremi di ricursione). Interessanti lavori si devono a H. Friedman che riprende precedenti ricerche di Wagner e Strong sulle strutture uniformemente riflessive (risalenti agli anni sessanta) e ancora a Y. Moschovakis. Possibili applicazioni su queste indagini sono le cosiddette *strutture enumerate*, sul cui dominio si definisce una enumerazione effettiva e che erano già state introdotte da Malcev, D. Lacombe e R. Fraïssé con l'obiettivo di fornire una base all'*algebra computabile*, ossia allo studio delle strutture algebriche definite in modo effettivo. In questa direzione già dagli anni cinquanta esistevano lavori assai interessanti di J. Shepherdson e M. Rabin che intendevano riprendere il vecchio progetto di un'algebra costruttiva risalente a

Kronecker e sviluppato poi parzialmente da B. van der Waerden.
Infine, anche se con maggior discontinuità, sono state approfondite le tematiche relative alla presentazione di «macchine» o comunque «dispositivi» in grado di simulare la generalizzazione della ORT a strutture arbitrarie (purché ovviamente dotate di una qualche possibilità di approccio di tipo combinatorio). Già Kleene nel suo studio sui funzionali ricorsivi aveva introdotto macchine generalizzate e nel periodo che stiamo esaminando si erano occupati di questo problema ancora Shepherdson e Friedman al primo dei quali – non va dimenticato – si deve l'introduzione di particolari macchine (note oggi come *macchine di Shepherdson-Sturgis*) che danno una formulazione estremamente maneggevole del concetto di computabilità, molto più duttile di quella data dalle macchine di Turing.

Come il lettore noterà ci siamo limitati a descrizioni schematiche senza entrare in dettagli. Ci sembra però che dovrebbe risultare chiaro il significato generale di queste indagini: non si tratta di generalizzare «a vuoto», ma di indagare il ruolo che le diverse componenti del concetto di effettivo (finitezza, computazione, procedura induttiva, ecc.) hanno e che spesso è difficile separare quando ci limitiamo al dominio dei naturali. L'obiettivo da una parte è certamente teorico e ha a che fare con problemi squisitamente logici, dall'altra ha un risvolto anche applicativo. Non è detto infatti che la precisazione del concetto generale di algoritmo fornita dalla ORT, e che riguarda essenzialmente i numeri naturali, sia il contesto adeguato per studiare gli algoritmi che si presentano in contesti specifici, per esempio in algebra, né sul piano dell'efficienza né su quello della comprensione concettuale. Una delle possibilità che la generalizzazione della ricorsività può offrire è uno sfondo concettuale in cui affrontare simili questioni.

1.3 *La teoria dei combinatori e il concetto di funzionalità*

Come abbiamo visto, la teoria della ricorsività e le sue generalizzazioni hanno come sfondo la teoria degli insiemi e in questo senso costituiscono un approccio *classico* al problema dell'effettività e della computabilità di funzioni e funzionali. Pur occupandosi di regole di computazione, la teoria della ricorsività studia infatti

funzioni, non *definizioni* di funzioni o regole di computo, e realizza così un approccio estensionale che si presenta come un capitolo della matematica classica su base insiemistica. Ciò è vero per tutte le impostazioni che abbiamo descritto. A partire dagli anni venti si era presentato però un altro contesto in cui sviluppare un'analisi diretta del concetto di computazione e di quello di regola: si tratta della teoria dei combinatori, che soprattutto negli ultimi decenni ha registrato notevoli sviluppi che hanno posto in luce un collegamento assai fecondo fra ricorsività, informatica e teoria delle categorie. È per questa ragione che ci sembra opportuno trattare ora l'argomento in un paragrafo a parte.

La ricchezza della problematica attuale legata alla teoria dei combinatori contrasta con le motivazioni apparentemente modeste che hanno portato al suo sorgere. L'idea di Moses Schönfinkel (cui si deve la creazione della teoria nel 1924) era di ridurre, all'interno del programma hilbertiano, i presupposti della sintassi dei linguaggi formali eliminando la distinzione in categorie sintattiche (variabili individuali, predicative, ecc.) che sembrava essere giustificata unicamente dall'interpretazione insiemistica che i linguaggi ammettevano, ma non dall'aspetto formale dei calcoli. L'idea di Schönfinkel era che in ogni calcolo logico si ha a che fare con termini che denotano funzioni, siano esse definite su individui o proposizioni, o che altro, e che le regole deduttive si possono vedere come prescrizioni riguardanti il *computo* di queste funzioni: sicché la sua analisi partiva appunto dal concetto di funzione o meglio – muovendosi sul piano sintattico – da quello di *termine* funzionale.

Per determinare una funzione occorre stabilire per ogni argomento per cui è definita il corrispondente valore. Pensiamo ad esempio alla sottrazione e supponiamo di disporre di un simbolo « $-$ » che indica l'operazione. Consideriamo l'espressione $x - y$: di per sé essa non denota una funzione ma la forma schematica del *valore* di una funzione per gli argomenti x e y. Per ottenere un termine che denoti la funzione occorre *astrarre* da $x - y$ rispetto alle variabili. Possiamo così convenire che il termine $\lambda xy.\ x - y$ denoti la *funzione* che applicata a ogni coppia a, b dà come risultato appunto $a - b$. Per ora non si tratta che di una semplice notazione, per distinguere la funzione in sé dal risultato della sua applicazione; perché l'interpretazione del termine $\lambda xy \cdot x - y$ sia quella voluta

occorre formulare regole per la sua computazione. Introduciamo allora tra i termini una operazione, quella di *applicazione*: dati due termini a e b, con (ab) indicheremo l'applicazione della «funzione» denotata da a all'«argomento» denotato da b. Per imporre questa interpretazione al termine $\lambda xy \,.\, x - y$ si dovrà allora formulare una regola (λ-*conversione*) del tipo

$$(\lambda x.b(x)a) = b[a]$$

dove $b[a]$ indica il risultato della sostituzione delle occorrenze *libere* di x (cioè non all'interno del campo d'azione di λ) con occorrenze di a. Nel nostro caso ad esempio applicando due volte la regola otterremo $((\lambda xy \,.\, x - ya)b) = ((\lambda y \,.\, a - y)b) = a - b$. In generale l'operazione di applicazione non è associativa e quindi per indicare l'ordine di applicazione dovremo usare delle parentesi; nell'esposizione informale che segue noi comunque le elimineremo ogniqualvolta ciò non si presti ad ambiguità.

L'analisi che la λ-regola ci offre del computo delle funzioni riduce sostanzialmente ogni funzione ad una unaria, con un solo argomento: la sottrazione tra due numeri si ottiene dal termine $\lambda xy \,.\, x - y$ mediante *due* successive applicazioni dell'operazione di astrazione λ. Da una prospettiva più generale possiamo vedere la λ-regola come una sorta di assioma di comprensione: se ammettiamo come esistenti nel nostro universo solo gli oggetti denotati dai termini, la regola ci dice infatti, poste delle condizioni su una funzione, che *esiste* la funzione in questione, così ad esempio il termine $\lambda xy \,.\, x - y$ denota la funzione richiesta. Ora quello che Schönfinkel voleva ottenere, a partire da uno stock finito di termini e dall'operazione di applicazione fra di essi, era la possibilità, per ogni termine X e variabili distinte $x_1,\dots,x_n$, di ottenere un termine F senza $x_1,\dots,x_n$ tale che

$$Fx_1,\dots,x_n = X$$

simulando cioè le proprietà di λ, così da giungere all'eliminazione delle variabili, in quanto il loro ruolo, come si è ben visto utilizzando l'operatore λ, è puramente accessorio, serve solo ad indicare in che senso si astrae da un termine dato, mentre l'essenziale è la sola regola su λ. Il problema era allora di trovare termini e re-

gole per il loro impiego che permettessero di «combinare» le espressioni richieste. Sono questi i *combinatori* definiti da Schönfinkel. Accanto all'eliminazione della categoria sintattica delle variabili, ne viene implicitamente fatta una seconda e più radicale, in quanto non si pone alcuna differenza fra termini che denotano funzioni e termini che denotano argomenti (come si può vedere tanto nel caso di λ che in quello di *F*).

In Schönfinkel il progetto rimase allo stato di abbozzo, e le sue motivazioni, che erano sostanzialmente riduzionistiche, non sembravano promettere sviluppi particolarmente interessanti. Questi temi vennero ripresi in una prospettiva più ampia nel 1932 e '33 da A. Church in due articoli dal titolo *A Set of Postulates for the Foundation of Logic I, II* (*Un insieme di postulati per i fondamenti della logica*), in cui appunto viene introdotto l'operatore λ che Church chiama *operatore di astrazione*, e utilizzato per formulare – sono parole di Church – «un sistema di logica matematica privo delle complicazioni della teoria dei tipi di Russell ma in grado d'altro canto di evitare i ben noti paradossi». Il sistema di Church – che si presentava come una versione della Grande Logica di tradizione fregeana in cui ricostruire l'intera matematica – oltre all'operatore λ utilizzava un operatore di descrizione ι (analogo a quello di Peano-Russell) un operatore di astrazione A, un operatore di generalità ristretta Π, l'esistenziale Σ, la congiunzione ∧ e la negazione ¬. Non importa in questa sede descrivere ulteriormente il sistema; basti dire che nel 1935 Kleene e Rosser ne dimostrarono l'incoerenza ricostruendo al suo interno il paradosso di Richard. L'importante dal nostro punto di vista è che qui compariva per la prima volta l'operatore λ e l'idea di sviluppare una teoria generale delle funzioni senza distinzioni di tipo entro la quale le funzioni venivano date in termini di regole di computo, di conversione, come ad esempio le due regole fondamentali

λ_1 $\quad$ $(\lambda x.\ a(x))b = a[b]$

λ_2 $\quad$ $\lambda x a(x) = a$ $\quad$ se x non è libera in a.

Queste regole riguardavano non oggetti astratti, ma *termini* di un linguaggio, cioè oggetti sintattici. In questo senso le funzioni venivano identificate con regole di computo.

Ritorneremo più avanti su questo aspetto, ma l'essenziale da

osservare è che non si trattava più di una semplice riduzione dell'armamentario linguistico, ma di una vera e propria *teoria*, che rifiutando gerarchie di tipi fissate *a priori* presentava le funzioni come tradizionalmente venivano intese, cioè come regole di calcolo, ponendo in primo piano l'aspetto intensionale (che riguarda la definizione) piuttosto che quello insiemistico estensionale. Resosi conto che le contraddizioni non riguardavano la teoria pura della λ-conversione, Church nel 1941 nel suo libro *Calculi of λ-conversion* (*Calcolo della λ-conversione*), isolava questa teoria delle funzioni e la presentava in sé e per sé. Già nel 1943 però Kleene, seguendo il suggerimento di Church di sviluppare la teoria dei numeri naturali in termini di λ-calcolo, aveva fatto un passo decisivo: identificato ogni naturale n con un termine $\bar{n}$ che applicato ad un termine funzionale f ci dà la n-esima iterazione della funzione, cosicché

$$\bar{n}fx = \underbrace{f(f,\ldots,f(x))}_{n\text{-volte}},$$

definiva direttamente la nozione di computabilità in termini di λ-calcolo. Nasceva così la cosiddetta Church-Kleene-computabilità, che lo stesso Kleene dimostrerà coincidere con la computabilità nel senso di Turing, di Herbrand-Gödel, ecc. come si è visto ampiamente nel capitolo III, 6.

Al di là delle non poche complicazioni formali, questo approccio risultava estremamente naturale da almeno due punti di vista: in primo luogo, nell'ottica del λ-calcolo ogni funzione è già in se stessa essenzialmente una regola di computo; in secondo luogo – ricordiamo – termini del linguaggio sono anche quelli del tipo XX in cui la «funzione» X si applica a se stessa. Questi termini ci permettono di catturare uno degli aspetti centrali della nozione di computabilità, il fatto cioè che definendo una funzione per dati argomenti è possibile sfruttare quel che sappiamo sulla funzione applicata ad altri argomenti, come ad esempio nelle definizioni ordinarie per ricorsione sui naturali, $f(n)$ può essere definita ricorrendo a suoi valori per $n' < n$. Principi come questi si chiamano in generale *principi di ricorsione* e stabiliscono schemi generali assai fecondi per definire funzioni computabili. I teoremi che stabiliscono la correttezza di un dato schema di definizione per ricorsione hanno in generale la forma di principi del *punto fisso* e coinvolgono, in

un modo o nell'altro, forme di autoapplicazione. Questo fatto su cui torneremo più avanti mostra come l'idea di funzione codificata nei λ-calcoli, legata com'è al concetto di regola, se da una parte fornisce un approccio diretto ai principi di ricorsione, dall'altra non è inseribile entro uno schema insiemistico in cui le funzioni sono insiemi di coppie e quindi, per l'assioma di fondazione, non si potrà mai avere che

$$f \in f$$

e cioè che f si applica a se stessa.

Questo spiega tra l'altro il carattere essenzialmente sintattico e formale che per lungo tempo ebbero le indagini sui λ-calcoli e sui combinatori che li estraniò per lungo tempo dal resto della ricerca logica. Non sembravano possibili infatti approcci di tipo semantico, in quanto la stessa nozione di semantica era legata a quella di insieme. La situazione è cambiata con lo svilupparsi della logica categoriale. In questo contesto è stato possibile fornire semantiche per la teoria dei combinatori generalizzando fondamentali ricerche iniziate da Scott al riguardo. Su un versante più legato alla teoria degli insiemi vanno invece ricordate le recentissime ricerche dovute soprattutto a Peter Aczel sugli insiemi non fondati, e il tentativo di fornire un contesto unitario per una teoria generale dei processi, che coinvolge anche le indagini informatiche.

Abbiamo parlato sempre dei λ-calcoli e di fatto c'è una sostanziale intertraducibilità tra λ-calcoli e combinatori nel senso di Schönfinkel come Rosser dimostrò nel 1935 nel suo lavoro *A mathematical Logic without Variables* (*Una logica matematica senza variabili*). In questo articolo Rosser costruiva una teoria sintetica dei combinatori equivalente a quella della λ-conversione di Church partendo da alcuni lavori precedenti di Haskell B. Curry che risalivano agli anni venti. È proprio Curry che in certo senso va considerato il padre della teoria dei combinatori, in quanto in un lungo giro d'anni sviluppò sistemi sempre più forti nel tentativo di fornire una fondazione per la matematica alternativa a quella insiemistica in cui non si assumessero gerarchie di tipo stabilite *a priori* e fossero possibili restrizioni di tipo finitista. Già nei primi anni venti Curry aveva costruito un *sistema basico dei combinatori*, estendendo il quale è possibile ottenere quella formulazione dei vari calcoli lo-

gici senza variabili e senza distinzioni *a priori* di categorie sintattiche che era l'obiettivo di Schönfinkel.

Una formulazione equivalente a quella di Curry è la seguente. Il linguaggio è formato da atomi, fra i quali tre combinatori fondamentali I, K, S. Termini saranno gli atomi e, se X e Y sono termini, anche (XY) lo sarà. Le regole daranno una relazione di *riducibilità* fra termini che intuitivamente, una volta che si interpreti (XY) come l'applicazione di X a Y, forniscono *regole di computazione*. Avremo allora gli assiomi (si legga $A \triangleright B$ come «A si riduce a B») per ogni X, Y, Z:

$$\mathrm{I}X \triangleright X$$
$$\mathrm{K}XY \triangleright X$$
$$\mathrm{S}XYZ \triangleright XZ(YZ)$$
$$X \triangleright X$$

e le regole:

$$(\alpha) \quad X \triangleright X' \Rightarrow ZX \triangleright ZX'$$
$$(\beta) \quad X \triangleright X' \Rightarrow XZ \triangleright X'Z$$
$$(\gamma) \quad X \triangleright Y \wedge Y \triangleright Z \Rightarrow X \triangleright Z.$$

Come dimostrato da Curry è possibile, sulla base di questo sistema, per ogni $x_1,\ldots,x_n$ e ogni M definire un termine (un combinatore) $[x_1,\ldots,x_n]$ che svolge il ruolo dell'operatore di astrazione λ, nel senso che per ogni termine $N_1,\ldots, N_n$ si ha

$$([x_1, \ldots, x_n]\ M)\ N_1, \ldots, N_n \triangleright [N_1/x_1, \ldots, N_n/x_n]\ M.$$

È questo il teorema di *completezza combinatoria* per il sistema basico che permette di stabilire l'equivalenza tra λ-calcolo e calcolo dei combinatori. Possiamo ora introdurre una distinzione tra termini: alcuni si possono contrarre, mediante gli assiomi e le regole, in altri termini, altri no. Più precisamente, se chiamiamo *redex* i termini della forma IX, SXY, e KXYZ, *contrarre* significa rimpiazzare in un termine l'occorrenza di un suo redex con il termine cui questo si riduce mediante regole e assiomi. Un termine è in *forma normale* quando è irriducibile, cioè quando non contiene alcun redex.

Questa distinzione ha un significato dal punto di vista interpretativo. Intuitivamente tutti i combinatori si possono vedere come

funzioni generalizzate in quanto fra di essi è definita un'operazione di applicazione; possiamo dire che gli argomenti su cui è definito un combinatore U sono quei termini X tali che UX si converte ad un termine normale, cioè irriducibile dal punto di vista della computazione. Non tutti gli UX però si convertono in forma normale e ciò si può interpretare quindi dicendo che *U non è definito* per X. Potremo quindi parlare di funzioni generalizzate, ma *parziali*, ossia non ovunque definite. In questa prospettiva divengono essenziali due questioni: in primo luogo stabilire un criterio affinché un termine U sia riducibile a forma normale; in secondo luogo, perché si possa davvero parlare di *funzioni* parziali, si dovrà dimostrare che non è possibile che un termine si converta a due termini in forma normale essenzialmente diversi.

Alla seconda questione rispondeva un teorema dimostrato da Church e Rosser negli anni trenta che avrebbe avuto un ruolo fondamentale nella teoria dei combinatori. Esso ammette due versioni: la prima afferma che se U si riduce a X e a Y, allora esiste uno Z tale che tanto X quanto Y si riducono a Z, dal che consegue che un termine si potrà convertire ad una sola forma normale. La seconda versione riguarda la relazione di equivalenza «=» che possiamo stabilire fra termini sulla base del concetto di riducibilità. La riducibilità infatti introduce chiaramente un preordine fra i termini: la «=» non è nient'altro che l'equivalenza che possiamo definire assumendo che $X = Y$ se e solo se $X \triangleright Y$ e $Y \triangleright X$. Il teorema di Church e Rosser afferma allora che se $X = Y$ esiste uno Z cui entrambi si riducono; da ciò consegue che se $X = Y$ o entrambi non hanno forma normale oppure, se l'hanno, è la stessa.

Rimane il primo problema, quello di determinare quando un termine U si converte a una forma irriducibile. Mentre esistono algoritmi per stabilire quando U e V hanno la stessa forma normale (ammesso che ne abbiano almeno una) non esiste in generale una procedura di decisione che ci dia tutti e soli i termini che ammettono forma normale.

Il sistema basico dei combinatori permette di costruire termini estremamente semplici; è però possibile estenderlo in modo da ricostruire sistemi formali più complessi, ad esempio fondare l'intera logica su questo sistema esteso. Si giunge così ai cosiddetti *sistemi illativi* di Curry che consentono la riformulazione dei calcoli lo-

gici fondamentali. L'idea di fondo è di ammettere fra gli enunciati della teoria anche enunciati della forma $\vdash X$ dove X è un termine, e che intuitivamente rappresentano l'asserto metateorico «l'enunciato X è dimostrabile». È questo un modo per passare dalla *computazione* di funzioni proposizionali alla *dimostrazione* di proposizioni. Si tratta poi di introdurre nuove combinazioni e quindi nuove costanti con cui formare le ordinarie connessioni fra enunciati: l'obiettivo finale è chiaramente quello di poter ottenere come dimostrabili nel nostro sistema tutte le espressioni della forma $\vdash X$ dove X è un termine che rappresenta una formula dimostrabile nel sistema logico che vogliamo rappresentare. Già a questo livello, come posto in luce da Curry, si poneva un grosso problema: la formulazione in termini di combinatori è estremamente *generale* in quanto le regole di conversione e le regole aggiuntive che dobbiamo dare per connettivi e quantificatori si applicano a termini qualsiasi, non avendo assunto alcuna distinzione in categorie sintattiche. Capita così che se assumiamo una regola del tutto naturale del tipo $X = Y \wedge \vdash X \Rightarrow \vdash Y$ e per ogni coppia di termini X e Y assumiamo altrettanto naturalmente che esista il termine $X \supset Y$, in modo che risultino soddisfatte le regole

$$[X \supset Y, X \vdash Y \text{ e } \vdash (X \supset (X \supset Y)) \supset (X \supset Y)]$$

otterremo che per ogni termine Y si deduce nella teoria $\vdash Y$; in altre parole, il sistema è contraddittorio. Occorreva quindi in qualche modo restringere la generalità delle regole e dopo molteplici tentativi Curry, con la collaborazione di Kleene e Rosser, riuscì a giungere a formulazioni adeguate.

Qui è importante non tanto dare una formulazione di questi sistemi quanto sottolineare l'atteggiamento di Curry di fronte a queste contraddizioni: in netta antitesi con l'approccio insiemistico che in base a un'analisi semantica assume una classificazione *a priori* dei termini in categorie sintattiche, per Curry si partiva invece dal sistema basico in cui tale distinzione non esiste e, man mano che si rafforzava il sistema aggiungendo nuovi operatori e nuove regole, si doveva analizzare il modo di imporre restrizioni alla generalità originale delle regole. La distinzione in categorie o in tipi ha un grosso ruolo in queste restrizioni, in quanto ci permette di porre distinzioni, in termini di tipi, all'applicabilità delle regole, *ma non viene accettata a priori e non è l'unica*: di volta in volta se

ne può riscontrare la necessità all'interno del sistema che si costruisce.

Ma qual è il vantaggio di questa tattica di partire da un livello così primitivo di generalità, a parte il suo interesse da un punto di vista strettamente teorico? Il fatto più significativo è che in questo modo abbiamo un contesto in cui studiare le regole di computo o di trasformazione tra oggetti nella loro generalità. Abbiamo sopra interpretato il sistema basico come una teoria del computo delle funzioni parziali. A ben vedere, come già anticipato, gli oggetti propri della teoria non sono le funzioni come oggetti insiemistici ma le *definizioni* di funzioni, le *funzioni in intensione*, come le chiamava Church. Il comportamento delle funzioni, in questo contesto, è dato unicamente dalle regole di conversione e i rapporti di uguaglianza o diversità tra termini valgono solo se si possono stabilire in base a queste regole. Ciò significa che la nostra è una teoria delle «funzioni parziali» *intensionali*, viste come procedure di computo e non come oggetti dell'universo degli insiemi. Questo spiega anche la non assunzione *a priori* di una gerarchizzazione in tipi, che si impone se si vedono le funzioni come insiemi di n-uple di oggetti (che saranno di tipo inferiore).

All'interno di questo contesto generale possiamo ricostruire teorie particolari riguardanti specifiche classi di funzioni. In particolare – come già abbiamo ricordato – Kleene ha ricostruito la teoria delle funzioni ricorsive all'interno del sistema basico dei combinatori. Si è già detto come in questo sistema si possono definire dei combinatori che rappresentano la successione dei numeri naturali; quello che Kleene prova è che è possibile trovare, per ogni funzione ricorsiva parziale, un termine del calcolo che la definisce all'interno del sistema. In altre parole, data una funzione n-aria f è possibile trovare un termine X del sistema tale che per ogni n-upla $a_1, ..., a_n$ di numeri naturali abbiamo $[X \overline{a}_1, ..., \overline{a}_n] = \overline{f(a_1, ..., a_n)}$ dove $\overline{a}_i$ è il termine che rappresenta il numero naturale a_i e $\overline{f(a_1, ..., a_n)}$ il termine che rappresenta il risultato. Come corollario dell'indecidibilità del problema della caratterizzazione delle funzioni ricorsive totali (*halting problem*) abbiamo allora l'indecidibilità del problema della riduzione a forma normale di cui sopra si è parlato.

Tutto ciò si applica ovviamente anche al λ-calcolo. Fu con questo sistema che Church dette la prima dimostrazione del suo teo-

rema di indecidibilità del calcolo dei predicati. I rapporti fra λ-calcolo e combinatori portano a introdurre diverse nozioni di riducibilità, oltre a quella da noi definita, a seconda che si pongano o meno certe restrizioni sulla convertibilità dei termini. In particolare si può assumere come schema di conversione la cosiddetta η-regola (cioè la λ_2 di sopra) che introduce un'assunzione di *estensionalità*, permettendo di raggiungere l'estensionale a partire dal contesto intensionale generale. Fra i risultati interessanti sugli analoghi del teorema di Church-Rosser per queste nozioni, notevole importanza ha il teorema di Corrado Böhm, del 1968, sulle forme normali forti (normali rispetto a una particolare nozione di riduzione).

La possibilità di rappresentare le funzioni ricorsive entro il sistema basico non solo offre un'ulteriore giustificazione alla tesi di Church, ma consente di cogliere la struttura unitaria di vari principi di recursione. Usando i combinatori molti di questi schemi si ottengono come corollari di un risultato centrale che riguarda i punti fissi dei termini in questo sistema. Ciò spiega il ruolo importante che la formulazione in termini di combinatori ha avuto in tempi recenti nello studio delle generalizzazioni delle funzioni ricorsive, primi fra tutti i funzionali di tipo finito introdotti da Gödel.

Questo ci porta al nucleo centrale degli sviluppi più recenti nello studio dei combinatori, cioè alla *teoria della funzionalità* nelle varie formulazioni introdotte da Curry e dai suoi allievi. Abbiamo già detto come la teoria basica si possa vedere come una teoria rudimentale delle funzioni parziali. Di qui si può partire per studiare combinazioni più complesse di funzioni una volta che questa parzialità venga assunta come concetto primitivo, introducendo una gerarchizzazione in tipi dei termini. Ciò non costituisce un abbandono dell'originario programma, ma piuttosto un suo approfondimento, volto a mettere in luce il ruolo che la gerarchizzazione in tipi può avere nella formulazione di schemi generali per la definizione di funzioni.

Ci sono due vie per formulare la teoria della funzionalità: una ammette una moltiplicazione dei combinatori fondamentali (I, K, S) per ogni tipo, l'altra assume che i combinatori possano ricevere infinite assegnazioni di tipo e pone in primo piano il meccanismo di questa assegnazione. Si introduce così una distinzione fra termini: termini che denotano tipi e termini più propriamente denotanti funzioni che vengono costruiti nel solito modo. In più si as-

sume una costante relazionale binaria *F* sì da leggere *FaX* come «ad *X* è assegnabile il tipo *a*», dove *X* è un termine; si tratta poi di fornire regole su *F* che stabiliscano il meccanismo dell'assegnazione. Naturalmente esiste una corrispondenza fra le due formulazioni, ma l'importante è che mediante una scelta opportuna dei tipi fondamentali (ad esempio assumendo un tipo per i numeri naturali, una costante 0 per lo zero e una *s* per il successore) possiamo costruire in modo unitario una teoria dei funzionali di tipo finito a partire dalle funzioni numeriche. Accanto ai combinatori fondamentali d'altra parte possiamo assumere altri combinatori che ci permettono di introdurre schemi di definizioni per funzioni sempre più complesse.

È così che dopo i lavori di Andrzej A. Grzegorczyk nel 1964, Luis E. Sanchis nel 1967 e William W. Tait nel 1965 e '69 si è riusciti a formulare la teoria dei funzionali di Gödel in modo combinatorio e più in generale a definire nozioni di computabilità per funzionali di tipo finito arbitrario. Anche per questi sistemi estesi si pone naturalmente il problema della convertibilità a termini normali e della unicità di questa convertibilità. Si tratta allora di estendere i risultati di Church e Rosser; questo è possibile, ma mentre per il sistema basico si può dare una dimostrazione dell'esistenza di forme normali che sfrutta solo l'induzione sui numeri naturali, per questi sistemi estesi occorre postulare principi d'induzione più forti (nel caso dei funzionali di Gödel l'induzione fino a ε_0, come già ricordato). La cosa importante è che l'uso dei combinatori in questo contesto permette un'analisi unitaria e profonda dei principi di ricorsione utilizzabili nella definizione di funzioni di tipo superiore.

Conosciamo il ruolo che questi funzionali hanno nella teoria della dimostrazione dopo la dimostrazione di coerenza dell'aritmetica data da Gödel nel 1958. La cosa si può estendere a teorie più complesse quali l'Analisi intuizionista, quella classica, ecc., e in generale (come fatto da Tait) è possibile, a partire dalla teoria dei combinatori, formulare una teoria generale del ragionamento costruttivo.

Vedremo più avanti più in dettaglio questi contatti fra teoria della dimostrazione e combinatori. La individuazione di questi contatti pone in luce la fecondità dell'approccio originario di Curry, ma nello stesso tempo fa risaltare un aspetto insoddisfacente di

tutta la teoria, costituito dal suo carattere esclusivamente formale. Si è posto così ripetutamente il problema di fornire una giustificazione concettuale del sistema basico e delle sue estensioni esibendo delle interpretazioni precise che rendessero possibile la sua lettura come teoria delle funzioni parziali. Uno dei risultati più significativi a questo riguardo ottenuto negli anni settanta è dato dalle interpretazioni che Dana Scott ne ha fornito in termini di teoria dei reticoli e ricorrendo a particolari spazi topologici. L'idea di fondo sfrutta la possibilità di ordinare i termini combinatori visti come funzioni parziali in base ai rapporti di inclusione dei loro domini. L'ordine che si ottiene in questo modo ammette una strutturazione in termini di reticoli completi sicché ad esempio si può stabilire un preciso contatto tra i teoremi del punto fisso validi per reticoli completi e analoghi teoremi dimostrabili nel calcolo dei combinatori. L'interpretazione di Scott avviene all'interno di una teoria generale della computazione che sembra aver ricevuto negli ultimi tempi notevole interessamento da parte dei *computer scientists*, ma una trattazione unitaria più ricca e profonda di tali problemi si è resa possibile in questi ultimi anni, utilizzando la teoria delle categorie.

2. SVILUPPI NELL'INDAGINE METAMATEMATICA

Già parlando degli sviluppi delle ricerche più strettamente fondazionali abbiamo avuto modo di ricordare alcuni dei risultati più interessanti ottenuti negli ultimi anni in quell'area di studi che genericamente possiamo chiamare metamatematica. Accanto però alle ricerche che considerano specificamente teorie particolari, esiste tutto un lavoro dedicato alla metamatematica generale di cui teoria della dimostrazione e teoria dei modelli costituiscono le articolazioni principali. Dagli inizi degli anni sessanta i contatti fra i diversi aspetti dell'indagine metamatematica (teoria della dimostrazione, dei modelli, problemi di decisione, ecc.) tendono sempre più a concretizzarsi spesso in analisi che è difficile isolare senza perdere il senso stesso dei problemi trattati. È per questa ragione che abbiamo pensato di riunire in questo paragrafo diversi argomenti che riguardano tutti l'indagine metamatematica e che

mostrano lo stretto interagire di considerazioni costruttive e non costruttive, il collegamento esplicito tra la pratica matematica e l'indagine di carattere puramente fondazionale.

Cominceremo con l'affrontare il tema dell'analisi della matematica intuizionista.

2.1 *Costruzioni e dimostrazioni: la matematica intuizionista*

Come già ricordato a suo tempo la problematica dell'intuizionismo ha un carattere nettamente diverso da quella degli altri indirizzi, nel senso che l'obiettivo fondamentale non è di giungere a una giustificazione o a una ricostruzione della matematica classica, ma piuttosto di costruire una matematica *alternativa* i cui concetti base abbiano contenuto costruttivo e abbiamo già avuto occasione di ricordare come dal 1931 in poi, a partire dalla formalizzazione della logica intuizionista, abbia avuto inizio il programma di assiomatizzazione della matematica intuizionista. Questo programma ha portato alla codificazione di settori centrali della pratica matematica costruttiva. Alla formalizzazione dell'aritmetica sono seguite varie formulazioni dell'Analisi, la formalizzazione della teoria delle specie, l'assiomatizzazione della teoria delle definizioni induttive iterate, ecc. Particolarmente interessanti sono state le varie assiomatizzazioni delle successioni di scelta dovute a Kreisel e Troelstra nel 1970, quelle delle successioni anomiche, dovute ancora a Kreisel nel 1968 ed i lavori sempre iniziati da Kreisel e portati avanti in particolar modo da Nicholas Goodman su una teoria delle costruzioni. Punto culminante di queste ricerche si può considerare l'edificazione di un sistema intuizionista di teoria dei tipi che Per Martin-Löf è andato sviluppando a partire dai primi anni settanta, come contesto unitario in cui formalizzare la matematica intuizionista e su cui avremo occasione di ritornare in seguito.

L'importanza di queste assiomatizzazioni consiste soprattutto nel fatto che esse hanno permesso di utilizzare metodi *metamatematici* per saggiare le alternative che le intuizioni di Brouwer lasciavano aperte, in particolare riguardo ai rapporti tra principi come la tesi di Church, la tesi di Markov e i vari principi brouweriani di continuità. I due metodi di tipo costruttivo che in questo contesto

si sono mostrati particolarmente utili sono stati da una parte la realizzabilità di Kleene e dall'altra l'uso di funzionali di tipo finito lungo la linea iniziata da Gödel. Con queste tecniche è stato ad esempio possibile analizzare entro che limiti le successioni di scelta si possono considerare come pure figure del discorso, ossia come essenzialmente eliminabili, mediante definizioni contestuali nello studio dell'Analisi elementare. Il problema è particolarmente interessante in quanto alcuni dei principi di continuità qui utilizzati non risultano sempre plausibili per le successioni di scelta e si presenta quindi come naturale l'idea di giustificare la quantificazione su successioni di scelta mediante reinterpretazioni. Questo è uno dei problemi che Kreisel e Troelstra hanno affrontato (dopo analoghi tentativi di Kleene via realizzabilità).

Un altro contesto in cui tanto la realizzabilità quanto soprattutto le interpretazioni via funzionali sono risultate utili è la *costruttivizzazione* di settori della matematica classica. Al di là delle indagini squisitamente matematiche e che hanno conosciuto notevoli sviluppi in questo giro d'anni sulla possibilità di avere corrispettivi intuizionisti di teorie classiche, rimane il problema generale di trovare metodi per determinare quando un risultato dimostrabile classicamente all'interno di una teoria lo è anche intuizionisticamente. Già abbiamo visto a suo tempo come questo sia stato possibile utilizzando traduzioni e altro tipo di interpretazioni nel caso della logica. Il fatto interessante è che i metodi sopra citati hanno permesso di stabilire per specifiche formulazioni intuizioniste dell'Analisi elementare e teorie collegate, regole derivate che garantiscono principi di transfer dalla matematica classica a quella intuizionista e che ricorrendo ai funzionali si sfruttano aspetti specifici delle teorie in esame, diversamente dai risultati logici generali tipo quelli di Glivenko.

In una prospettiva più autonoma rispetto all'intuizionismo ma certamente legata ad esso si situano le ricerche di Errett Bishop sull'Analisi costruttiva iniziate attorno al 1967 e che dal punto di vista strettamente matematico costituiscono l'evento più importante delle indagini sulla matematica costruttiva degli anni sessanta e settanta. Partendo da un programma in netto antagonismo con la fondazione insiemistica e con l'approccio hilbertiano, Bishop è riuscito a dimostrare come un'ampia parte dell'Analisi classica si possa ricostruire su basi costruttive. Non avrebbe senso

entrare nei dettagli della descrizione di tutti questi sistemi; quello che per noi conta è che essi testimoniano, con le discussioni che hanno suscitato circa la loro accettabilità dal punto di vista costruttivo, la fecondità dei postulati iniziali del costruttivismo e la possibilità di concretare i suoi programmi in una vera e propria nuova pratica matematica. Diversamente da Brouwer gli interessi di Bishop non sono direttamente fondazionali. Pur avendo un glorioso passato di analista classico alle spalle, Bishop è infatti convinto che la matematica astratta di tipo insiemistico (non esclusi i tentativi eterodossi come l'Analisi non standard) testimonino un progressivo allontanamento dal contenuto reale della matematica e si riducano spesso a formalismi giustificati solo sul piano linguistico. Per Bishop la matematica deve essere in ultima istanza trascrivibile in un linguaggio interpretabile direttamente in senso numerico, il che comporta da una parte il rigetto di alcuni sviluppi estremi tipo ad esempio la teoria delle successioni di libera scelta, dall'altra l'abbandono di alcuni principi della logica classica, quale il terzo escluso. Bishop non ha mai «formalizzato» la sua concezione della matematica e questo è stato tentato da altri, ad esempio Nicolas Goodman e Feferman che hanno costruito sistemi entro cui sviluppare una matematica costruttivista indagandone le proprietà metateoriche.

Differenti sono i problemi che si pongono dal punto di vista più strettamente fondazionale. Per accertare le potenzialità dei sistemi costruttivisti, per determinare qual è lo *status* di principi proposti e discussi quali ad esempio lo schema di Markov o altri principi brouweriani, si è imposta – come detto – un'analisi metamatematica di questi sistemi. Dal nostro punto di vista le novità più interessanti delle indagini sull'intuizionismo a partire dagli anni sessanta riguardano appunto questo aspetto. Come nel caso classico, lo studio dei sistemi intuizionisti si può condurre sia dal punto di vista della teoria della dimostrazione (analisi della struttura delle prove) sia da quello semantico (mediante l'elaborazione di interpretazioni delle teorie). È in questa seconda direzione che l'indagine si stacca decisamente da quella analoga sui sistemi classici. Se infatti è possibile – come vedremo – fornire interpretazioni dei sistemi intuizionisti sulla base della teoria degli insiemi (paradigmi di queste interpretazioni sono l'analisi in termini di strutture di Kripke o di strutture topologiche) si pone d'altra par-

te il problema ben più urgente dell'elaborazione di interpretazioni, modelli plausibili dal punto di vista costruttivo e quindi di necessità non insiemistiche: è attraverso queste infatti che è possibile non solo valutare la potenza delle teorie in esame (cosa questa possibile anche con costruzioni classiche) ma per così dire estrarne il contenuto costruttivo che costituisce la loro ragion d'essere.

Tentativi di formulare interpretazioni di questo tipo si erano succeduti fin dagli anni quaranta, e i due esempi paradigmatici sono sostanzialmente la realizzabilità di Kleene con le sue molteplici diramazioni e l'interpretazione mediante funzionali proposta da Gödel nel 1958. Partendo dall'idea base che ogni asserto intuizionista è un asserto sull'esistenza di una costruzione, tutte queste proposte hanno avuto come scopo quello di individuare la classe delle operazioni costruttive che garantiscono la validità degli asserti appartenenti a date teorie intuizioniste. In sostanza tutte queste interpretazioni si basano sull'analisi delle formule della teoria formale mediante funzionali che estraggono il loro contenuto costruttivo: ad ogni formula $\mathscr{A}$ della teoria viene associata una formula $\mathscr{A}^*$ del sistema interpretante che «traduce» in termini di esistenza di funzionali il significato della formula $\mathscr{A}$.

Il carattere costruttivo di questa interpretazione può essere giustificato in varie maniere e nel 1967 Tait ha proposto, partendo proprio dal sistema T dei funzionali di Gödel, un modo estremamente fecondo per far questo. L'idea, che si riconoscerà come familiare dopo aver letto le pagine sui combinatori, è che il sistema e gli schemi definitori per i funzionali primitivi ricorsivi vanno visti come *regole di computazione* per termini piuttosto che come schemi definitori. La differenza è significativa in quanto in questo modo si dà alla teoria di Gödel un'interpretazione intensionale e non estensionale: gli oggetti in esame non sono più entità astratte insiemistiche (funzionali) ma *definizioni* di funzionali cioè funzionali in intensione. Sono questi i veri e propri oggetti costruttivi. Il problema centrale è quello di determinare quando valgono certe equazioni tra termini che denotano funzionali. Se, come proposto da Tait, accettiamo il punto di vista intensionale, queste equazioni non vanno più riguardate come identità fra gli oggetti designati dai termini, bensì come asserti che stabiliscono che le procedure di computazione associate ai singoli termini coincidono; il problema è allora quello di dimostrare queste coincidenze sulla base del-

le regole di computo che definiscono i termini introdotti: alla identità tra funzioni si sostituisce dunque l'eguaglianza definitoria (*definitional equality*). Il modo più adeguato per analizzare da questo punto di vista i funzionali è quello di introdurli secondo lo stile della logica combinatoria, ché allora diviene possibile determinare l'eguaglianza definitoria tra due termini t e t' che «denotano funzionali come convertibilità di t e t' ad uno stesso termine t''» in forma normale. La teoria dei combinatori si rivelava così come il contesto più naturale per lo studio delle funzioni dal punto di vista intensionale o in altre parole per lo studio delle procedure di computo dei termini.[3]

La valutazione del carattere costruttivo della classe di funzionali introdotta dipende dalla plausibilità delle procedure induttive necessarie o in altre parole dalle proprietà delle regole di computazione. In questo contesto è possibile reintrodurre l'approccio estensionale postulando, oltre alle ordinarie regole di conversione, il già citato schema η secondo il quale, se x non è libera in b,

$$(\eta) \qquad \lambda x . b(x) \text{ si converte in } b,$$

che non è altro che una forma di assioma di estensionalità. Dal punto di vista costruttivo il problema è se l'assunzione di questo schema sia corretta e su questo punto le discussioni sono state intense. Comunque, quale che sia la posizione sull'estensionalità, una volta che si sia dimostrata l'esistenza e l'unicità di forme normali è possibile costruire un modello di termini per la teoria in esame. Modello del sistema T di Gödel ad esempio sarà l'universo delle classi di equivalenza dei termini rispetto all'eguaglianza definitoria, su cui è definita l'operazione di applicazione determinata dalle regole di computazione. È in questo senso che l'approccio in termini di computabilità ci fornisce *modelli* per le teorie esaminate.

Dal lavoro di Tait in poi questo approccio alla interpretazione delle teorie intuizioniste si è sempre più esteso e ha portato a risul-

[3] Si vede anche come il problema centrale di provare che quelle date sono definizioni di funzioni si traduca nel seguente: dimostrare che per ogni termine esiste una sua forma normale (esistenza) e che questa forma è essenzialmente unica (teorema di Church-Rosser). Mentre l'unicità per il sistema T di Gödel si può dimostrare con mezzi strettamente finitisti, in generale l'esistenza comporta il ricorso a procedure che possono essere non finitiste.

tati decisamente interessanti. Basti pensare ai lavori dello stesso Tait sui funzionali a recursione sbarrata (1971) necessari per interpretare l'Analisi intuizionista, e ai lavori del 1970 di W.A. Howard sul collegamento fra computazione di termini e induzione su ordinali in cui viene analizzato in dettaglio come la dimostrazione dell'esistenza di una forma normale per ogni termine che definisce funzionali ricorsivi primitivi sia legata all'assegnazione di ordinali a questi termini e all'assunzione di principi di induzione sul segmento di ordinali utilizzato.

In questa prospettiva di ricerca di un concetto di modello per teorie intuizioniste si situa un altro lavoro di Martin-Löf del 1975, *About models for intuitionistic type theories and the notion of definitional equality* (*Sui modelli delle teorie intuizioniste dei tipi e la nozione di uguaglianza definizionale*) che formula una nozione generale di modello per una classe molto ampia di queste teorie: escluse sono quelle che contengono assiomi per successioni di scelta. Martin-Löf insiste sulla necessità di *escludere* un'interpretazione estensionale dell'eguaglianza come non adeguata dal punto di vista costruttivo e propone la sua sostituzione, nella costruzione di modelli intuizionisti, con l'eguaglianza definitoria. Questo porta al rifiuto della regola η e quindi rende impossibile provare la validità di certi principi rispetto alle interpretazioni su modelli di termini in cui la convertibilità è così ristretta. L'idea di Martin-Löf è che per avere senso costruttivo una relazione di eguaglianza deve essere decidibile e se si assume come base per la dimostrazione di questa decidibilità quella della relazione di conversione, non si può in alcun modo accettare che l'identità fra termini coincida con l'identità fra gli oggetti denotati: la convertibilità ci dà infatti un'eguaglianza intensionale che non si può confondere con quella fra gli oggetti astratti denotati. Per questa ragione, secondo Martin-Löf, se vogliamo avere come modello quello dei termini, non si può accettare la validità dell'assioma $(a = b) \vee \neg (a = b)$ per gli oggetti di tipo superiore. È notevole il fatto che entro lo schema generale di Martin-Löf con questa vincolante è possibile riottenere le varie interpretazioni in termini di realizzabilità proposte a partire da Kleene.

Se passiamo ora alla teoria della dimostrazione, troviamo che anche in questo contesto è essenziale il ruolo di una formulazione in termini di combinatori e qui non si verifica quel divario netto

fra teoria della dimostrazione e teoria dei modelli che esiste nel caso classico. C'è infatti uno strettissimo legame fra l'elaborazione di modelli di termini come concepita sopra e l'analisi delle dimostrazioni di una teoria formulata stile deduzione naturale. La possibilità stessa infatti di dimostrare che una certa struttura di termini è modello di una teoria è strettamente legata alla possibilità di fornire un'analisi combinatoria (nel senso della teoria dei combinatori) della struttura delle dimostrazioni della teoria: in quest'ottica, le regole di deduzione ci danno un modo per determinare, date le costruzioni per le premesse, le costruzioni per le conseguenze. È quello che si verifica nel caso classico dimostrando il teorema di validità. Il rapporto è duplice: da una parte va nel senso detto sopra cioè ci permette di dimostrare che certi insiemi di termini sono modelli, dall'altra ci consente di dimostrare teoremi di normalizzazione per le dimostrazioni di una teoria.

Quello che sta alla base di questa possibilità è il collegamento fra calcoli della deduzione naturale e teoria della funzionalità cui si è accennato sopra parlando della teoria dei combinatori. Questo collegamento era reso possibile dal nuovo approccio ai calcoli della deduzione naturale presentato da Dag Prawitz a partire dal 1965, all'interno del suo progetto di una teoria non riduttiva della dimostrazione, da noi trattato nel capitolo precedente. Il metodo risaliva a Curry e successivamente era stato sviluppato da William Howard. Diamo un'idea di cosa questo possa significare in un caso particolarmente semplice. Supponiamo di avere un linguaggio del primo ordine i cui unici operatori logici siano $\supset$ e $\forall$. Ad ogni formula $\mathscr{A}$ del linguaggio associamo un'infinità di variabili $x_i^{\mathscr{A}}$ e intuitivamente interpretiamo $x_i^{\mathscr{A}}$ come l'i-esima assunzione di una dimostrazione di $\mathscr{A}$. Useremo poi a, b, c, ... come nomi per dimostrazioni. Parallelamente introduciamo un insieme infinito di variabili x_i^I che intenderemo scorrere su nomi di individui; useremo poi t, s,... come nomi di individui. Scriviamo inoltre $a \models \mathscr{A}$ per dire che a è una dimostrazione di $\mathscr{A}$. A questo punto possiamo collegare a dimostrazioni nel calcolo della deduzione naturale termini costruiti a partire dalle variabili e le costanti sopra date utilizzando l'operatore di astrazione e l'operazione di applicazione. Per far questo basta correlare ad ogni applicazione di regole un termine opportuno. Cominciamo dalle regole di introduzione, che nel nostro caso saranno

$$I \supset \quad \dfrac{\begin{matrix}[\mathscr{A}]_i \\ \mathscr{B}\end{matrix}}{\mathscr{A} \supset \mathscr{B}} \qquad\qquad I\forall \quad \dfrac{\begin{matrix}\vdots \\ \mathscr{B}\end{matrix}}{\forall x_i^I \mathscr{B}}.$$

Ad una applicazione di $I \supset$, supponendo che $b \models \mathscr{B}$, possiamo associare il termine $\lambda x_i^{\mathscr{A}}. ab$. La cosa è intuitiva in quanto il termine viene a denotare quell'operazione che ad ogni assunzione $x_i^{\mathscr{A}}$ di $\mathscr{A}$ fa corrispondere una dimostrazione di $\mathscr{B}$; mediante l'applicazione di λ, $x_i^{\mathscr{A}}$ diviene vincolata e il significato della regola $I \supset$ è esattamente questo, di «scaricare» un'assunzione di $\mathscr{A}$. In generale le variabili libere in un termine indicano le assunzioni da cui la dimostrazione ad esso associata dipende. Discorso analogo per $I\forall$: all'applicazione della regola associeremo il termine $\lambda x_i^I. b$. Anche qui il termine denota quella funzione che applicata a ogni individuo t ci darà una dimostrazione di $\mathscr{B}(t)$. La restrizione sulla regola di conversione per $\lambda x_i^I. b$ corrisponde alla restrizione sull'applicazione della regola nel calcolo. Passiamo ora alle regole di eliminazione.

$$E \supset \quad \dfrac{\begin{matrix}\vdots & \vdots \\ \mathscr{A} \supset \mathscr{B} & \mathscr{A}\end{matrix}}{\mathscr{B}} \qquad\qquad E\forall \quad \dfrac{\forall x_i^I \mathscr{B} \quad t}{\mathscr{B}[t/x_i^I]}.$$

Qui l'operazione sui termini sarà l'applicazione. Precisamente faremo corrispondere a $E \supset$ un termine ca dove $c \models \mathscr{A} \supset \mathscr{B}$ e $a \models \mathscr{A}$. Intuitivamente: la dimostrazione di $\mathscr{A} \supset \mathscr{B}$ è una funzione che ad ogni dimostrazione di $\mathscr{A}$ associa una dimostrazione di $\mathscr{B}$. Analogamente a $E\forall$ associeremo il termine $c(t)$ dove t è il nome per un individuo e $c \models \mathscr{B}$: c è quindi una funzione che a ogni nome di individuo t associa una dimostrazione di $\mathscr{B}(t)$. Poiché ogni dimostrazione non è che una successione finita di applicazioni di queste quattro regole, ad ogni dimostrazione sarà univocamente associato un termine.

In questo modo si è data una formulazione precisa dell'interpretazione intuitiva di $\supset$ e di $\forall$ fornita da Heyting, quindi un'analisi del significato costruttivo degli operatori, che è parallela a quella fornita dalle regole. Quanto alle dimostrazioni delle eventuali assunzioni, la loro natura dipende ovviamente dal tipo di assunzioni con cui abbiamo a che fare e quindi dal tipo di teoria che

stiamo analizzando. Se esse sono equazioni, le dimostrazioni potranno essere prove della convertibilità ad una stessa forma normale di dati termini. Il collegamento si può estendere a linguaggi più ricchi di quello da noi esaminato; ciò che conta è che in questo modo è possibile stabilire una precisa corrispondenza fra i passi che ci possono permettere di ridurre una dimostrazione in forma normale e le regole di conversione per termini. Consideriamo ad esempio la dimostrazione

$$\dfrac{\dfrac{\begin{matrix}[\mathscr{A}]\\ \mathscr{B}\end{matrix}}{\mathscr{A}\supset\mathscr{B}} \quad \mathscr{A}}{\mathscr{B}} \qquad \text{che si riduce a} \qquad \dfrac{\begin{matrix}\vdots\\ \mathscr{A}\end{matrix}}{\mathscr{B}}$$

dove è stato eliminato il *detour* costituito dall'applicazione di una introduzione seguita da una eliminazione. Se consideriamo i termini corrispondenti avremo questa situazione. Alla prima dimostrazione corrisponde il termine $(\lambda x^{\mathscr{A}}.b)a$; alla seconda $b[a]$, vale a dire una dimostrazione di b che dipende da a (il termine che si ottiene da b che contiene una variabile libera sostituendola con a). Dalla regola della λ-conversione abbiamo che il primo termine si converte nel secondo, quindi al passo di riduzione corrisponde la conversione dei termini associati.

La cosa si può estendere a tutte le altre regole, utilizzando, anche, oltre alla λ-conversione, la regola η. Per sapere se una dimostrazione si converte in forma normale basta quindi sapere se il termine associato si converte a forma normale applicando le regole del calcolo dei combinatori. Questa corrispondenza si può modificare ed estendere a seconda della teoria e della logica in esame. Il fatto essenziale che questo isomorfismo tra funzioni e dimostrazioni comporta, è la possibilità di vedere le proposizioni come *tipi*, identificando una proposizione $\mathscr{A}$ con la collezione delle sue dimostrazioni. A questo punto le regole di formazione delle formule ci suggeriscono nuove operazioni sui tipi, ed è possibile una trattazione unitaria che ha un grande significato in quanto si realizza all'insegna dell'unico concetto di costruzione.

È proprio partendo da queste considerazioni che Martin-Löf ha sviluppato la sua teoria intuizionista dei tipi, in cui gli enti primitivi sono appunto i tipi visti come prescrizioni che ci dicono che

cosa bisogna fare per costruire un oggetto del tipo dato, e le funzioni, viste come regole. Le proposizioni vengono identificate, secondo l'interpretazione illustrata sopra, con le classi delle loro dimostrazioni che sono appunto funzioni appartenenti al tipo dato dalla proposizione.

A questo punto regole di formazione per proposizioni e per tipi coincidono, cosicché una stessa operazione sarà da una parte un operatore logico, dall'altra un'operazione su tipi. All'interno della teoria si assume anche l'esistenza di particolari tipi, che godono di specifiche proprietà di riflessione e che costituiscono gli *universi*, il cui ruolo è di permettere la definizione di concetti come quello di sequenza finita. Imponendo la chiusura rispetto a operazioni significative quali quella di prodotto cartesiano di famiglie arbitrarie (che ci garantisce in particolare l'esistenza dell'insieme di tutte le funzioni da un insieme di un tipo a un altro) è possibile per Martin-Löf sviluppare una teoria predicativa in cui però non si presentano le difficoltà tipiche delle teorie ramificate tradizionali e il cui linguaggio è estremamente più duttile e perspicuo. Non ultimo vantaggio è da una parte la sua possibilità di essere usato come linguaggio di programmazione, dall'altra l'essere in grado di fornire una trattazione adeguata di concetti in uso nella teoria delle categorie (quali quelli di funtore, trasformazione fra funtori, ecc.) che nella teoria degli insiemi classica risulta macchinoso trattare. Ulteriore prova di questo fatto è la stretta simmetria che esiste tra le regole di introduzione ed eliminazione associate alle operazioni su tipi e le aggiunzioni tra funtori che esse determinano in un contesto categoriale.

L'isomorfismo di cui sopra abbiamo parlato fra termini e dimostrazioni, con la conseguente intertraducibilità tra conversione e riducibilità tra dimostrazioni (nel senso introdotto da Prawitz) si è rivelato uno strumento fondamentale nella dimostrazione dei teoremi di normalizzazione che in questa prospettiva divengono teoremi di forma normale per *termini* costruiti in un linguaggio come quello dei combinatori con distinzioni di tipo. Collegando questo isomorfismo con le tecniche introdotte da Tait di cui abbiamo già parlato, Martin-Löf è riuscito nel 1970 a dimostrare un teorema di normalizzazione per la teoria delle definizioni induttive iterate e, sempre nello stesso anno, per la teoria delle specie, una volta formulate queste teorie in termini di calcoli della derivazione naturale alla Prawitz.

Per apprezzare il senso di questi risultati è opportuno tener presente quali sono le difficoltà che ogni dimostrazione di un teorema di normalizzazione deve affrontare in questo caso. Formulando la teoria delle specie, tra le regole compare quella di eliminazione del quantificatore universale

$$\frac{\forall X A(X)}{A(T)}$$

dove T è una formula. L'impredicatività della logica del secondo ordine (e la teoria delle specie è essenzialmente una logica del secondo ordine intuizionista) sta nel fatto che una formula del tipo $\forall X A(X)$ in una dimostrazione Σ può essere rimpiazzata da una nuova formula $A(T)$ in una dimostrazione Σ' cui Σ si riduce e $A(T)$ può essere notevolmente più complicata di $\forall X A(X)$, in particolare T può contenere $\forall X A(X)$ come parte. È difficile vedere quindi come si potrebbe provare che dopo un numero finito di riduzioni immediate si può arrivare ad una forma normale. Martin-Löf supera la difficoltà mediante il *metodo della computabilità* – è questa l'espressione che egli usa – mostrando che le tecniche di Tait di convertibilità si applicano anche alle dimostrazioni formali.

Più o meno negli stessi anni Prawitz usando tecniche simili provava un analogo teorema di normalizzazione per la logica *classica* del secondo ordine, giungendo così alla dimostrazione della congettura di Takeuti. I metodi di Prawitz, diversamente da quelli usati nelle precedenti dimostrazioni di cui abbiamo parlato a suo tempo, non sono semantici e hanno un carattere almeno parzialmente costruttivo, in quanto appunto egli non si limita a dimostrare l'esistenza di una forma normale, ma dimostra la *normalizzazione* di ogni dimostrazione del calcolo, vale a dire la sua trasformabilità mediante le regole di riduzione in una dimostrazione di forma normale. L'elemento di non costruttività sta nella forza dei principi induttivi che deve usare per ottenere questo risultato.

Sempre nei primi anni settanta il risultato forse più significativo a questo riguardo – da cui in particolare dipendono i risultati di Prawitz e Martin-Löf sulla teoria delle specie – è l'estensione dell'interpretazione funzionale di Gödel all'Analisi intuizionista data da Jean-Yves Girard. L'idea di fondo di Girard è di definire una nozione di validità per dimostrazioni (che egli chiama *candidates de*

reducibilité) che costituisce una sorta di riducibilità astratta fra termini o (il che è lo stesso) fra formule e che gli permette di neutralizzare l'impredicatività del tipo visto sopra. È questa idea di fondo che, in forme diverse, è stata utilizzata poi da Prawitz introducendo, in stretta analogia con i candidati di riducibilità, una nozione di validità per dimostrazioni rispetto a particolari *assegnazioni* ai termini del secondo ordine.

Diversamente da Prawitz, Girard utilizza l'idea di Gödel di interpretare l'Analisi all'interno di una teoria Y di funzionali di tipo finito. Caratteristica di questa teoria è che oltre a contenere nuovi tipi (tra i quali quelli che corrispondono alle operazioni logiche di implicazione e quantificazione) contiene anche *variabili* sui tipi, che possono occorrere nei termini, i quali possono così avere un tipo variabile. Girard riesce a interpretare l'Analisi e la teoria dei tipi entro il sistema Y e a provare così che se $\mathscr{A}$ è un teorema dell'Analisi la sua trasformata $\mathscr{A}^*$ è valida nel sistema Y. Utilizzando il collegamento formule-tipi, può dimostrare poi il teorema di normalizzazione per la logica intuizionista del secondo ordine G^1 LC (secondo la formulazione di Takeuti) e per la teoria dei tipi d'ordine finito G^ω LC. È difficile sottovalutare l'importanza dei risultati di Girard e soprattutto delle tecniche introdotte, che portano ad un netto passo avanti rispetto alle precedenti indagini sulla congettura di Takeuti per il loro carattere sintattico e almeno parzialmente costruttivo.

Questo carattere costruttivo è dovuto al fatto – come già abbiamo ricordato – che quello che si dimostra è un teorema di normalizzazione, non semplicemente di esistenza delle forme normali, cosicché, come Girard sottolinea, non utilizza il terzo escluso e risulta formalizzabile entro l'Analisi, rispettivamente, la teoria dei tipi, impredicative intuizioniste. Questa dimostrazione, come quelle di Prawitz e di Martin-Löf, non risulta allora utilizzabile nel contesto di una teoria riduttiva, in quanto la metateoria usata è forte almeno quanto la teoria esaminata, ma costituisce un tipico esempio di risultato proprio della teoria generale della dimostrazione e il suo interesse sta nel fatto che permette di ottenere come corollari informazioni importanti, quali ad esempio che tanto la teoria delle definizioni iterate quanto la teoria delle specie – come l'aritmetica intuizionista – godono della proprietà della disgiunzione e dell'esistenziale già ricordate.

Se questi risultati mostrano la possibilità di un approccio che, ponendo in primo piano aspetti intensionali, permette una fusione fra metodi sintattici e semantici nel contesto di un'ottica costruttivista, risultati interessanti sono stati ottenuti però anche in altri contesti utilizzando questa volta metodi francamente non costruttivi, più precisamente modificazioni o arricchimenti delle semantiche non costruttive della logica intuizionista. In una serie di lavori del 1968 Scott riusciva infatti ad estendere l'interpretazione topologica della logica intuizionista data da Tarski, all'Analisi intuizionista e alla logica del secondo ordine fornendo modelli estremamente interessanti che se pur non hanno in quanto tali una giustificazione costruttiva, lumeggiano le alternative che si presentano una volta che si vogliano chiarire formalmente le idee di Brouwer. In questa stessa direzione si situano alcuni lavori di C. Smorynski che estendono la semantica di Kripke al secondo ordine, e di Dov Gabbay che hanno gettato le basi per una vera e propria teoria dei modelli intuizionista fondata sulla semantica di Kripke. È risultato così possibile dimostrare analoghi dei risultati classici sulla completezza della teoria dei campi algebricamente chiusi e di quella dei campi reali chiusi, analoghi del teorema di interpolazione di Craig, ecc. L'interesse di queste ricerche sta in due fatti sostanzialmente: in primo luogo esse hanno permesso di ottenere informazioni sulla metamatematica dei sistemi intuizionisti che difficilmente si sarebbero potute ottenere utilizzando esclusivamente strumenti costruttivi; in secondo luogo hanno consentito di collegare la matematica intuizionista con alcuni aspetti fondamentali della ricerca classica. I collegamenti riguardano in particolare i modelli booleani e la rappresentazione di strutture algebriche, quali gli anelli o i moduli, come sezioni di fasci. In altre parole è possibile vedere certi aspetti dell'Analisi o dell'algebra intuizioniste in un contesto diverso da quello intuizionista «classico». Come dimostrato da Christopher H. Mulvey c'è uno stretto collegamento tra lo studio dei fasci di strutture algebriche e la teoria intuizionista di queste strutture; analogamente c'è un collegamento tra i modelli topologici dell'Analisi definiti da Scott e l'interpretazione dell'Analisi all'interno di un topos di prefasci o fasci.

Più avanti avremo occasione di tornare su alcuni di questi temi; vogliamo concludere invece accennando a un problema che riguarda i rapporti fra matematica classica e intuizionista in un'ac-

cezione diversa da quella di cui si è detto fino ad ora: entro che limiti le dimostrazioni di completezza per la logica intuizionista rispetto alle diverse semantiche (algebrica, topologica, di Kripke, di Beth, ecc.) sono accettabili dal punto di vista intuizionista? Già Beth, nella definizione dei suoi modelli, aveva individuato nelle successioni di scelta un ponte di collegamento fra le sue interpretazioni, che sono oggetti classici, e la matematica intuizionista. Negli anni cinquanta Kreisel assieme a Verena Dyson aveva approfondito questi legami, ma rimanevano dubbi – che già Gödel aveva espresso – sulla possibilità di dare versioni costruttive di questi risultati di completezza. Nel 1970 sempre Kreisel ha infatti dimostrato che assumendo la tesi di Church sotto condizioni assai plausibili, l'insieme delle formule «valide» della logica intuizionista non può essere r.e. e quindi che *nessuna* delle semantiche presentate è completa rispetto al calcolo di Heyting (il cui insieme di teoremi è ovviamente r.e.) e può considerarsi una semantica adeguata dal punto di vista intuizionista. In tempi più vicini a noi d'altra parte, W. Veldman e H.C.M. de Swart hanno cercato di fornire dimostrazioni intuizionisticamente plausibili di completezza rispetto a semantiche basate su quelle di Kripke e di Beth, ma non sembra ci sia accordo su questa plausibilità, come dimostrano i dibattiti su questo argomento. Per converso, sul piano classico, già dalla fine degli anni sessanta sempre più chiari sono risultati i rapporti fra le semantiche algebrica e topologica e quelle di Beth e di Kripke, la cui intertraducibilità discende da teoremi di rappresentazione per le algebre di Heyting su algebre di aperti di spazi topologici ottenuti mediante preordini (si veda il paragrafo 7 del capitolo II).

2.2 *Teoria della dimostrazione*

Lo studio della matematica costruttivista e in particolare intuizionista, di cui ci siamo appena occupati, è sicuramente stato negli anni sessanta, e continua ad essere tutt'oggi, uno dei campi privilegiati d'applicazione della teoria della dimostrazione e non a caso Kreisel, nel 1965, in un articolo panoramico dal titolo *Mathematical Logic* (*Logica matematica*) basava la sua esposizione degli sviluppi della teoria della dimostrazione esemplificando

proprio sui sistemi intuizionisti. Il raggio d'applicazione della teoria della dimostrazione è tuttavia ben più vasto e le direttive di ricerca fondamentali sono fino agli anni settanta quelle cui abbiamo accennato nel capitolo precedente. Non c'è possibile qui entrare in dettagli ulteriori su ricerche più particolari che abbiamo avuto occasione di ricordare, quali gli sviluppi delle indagini sui sottosistemi dell'Analisi, gli studi di Feferman e Kreisel sulle progressioni autonome, quelli di Schütte e della sua scuola sugli ordinali che intervengono nella dimostrazione dei teoremi tipo *Hauptsatz* o la riscoperta dei metodi di Herbrand per la coerenza verso la metà degli anni settanta.

Ci limiteremo piuttosto a sottolineare è il *mutamento di prospettiva* che si va consolidando in questo campo alla fine degli anni sessanta. Un'illustrazione concreta di ciò si può avere confrontando la prima e la seconda parte dell'articolo di Kreisel *A survey of proof theory* (*Panorama di teoria della dimostrazione*) uscite rispettivamente nel 1968 e nel 1971. Mentre la prima parte è essenzialmente dedicata alla teoria della dimostrazione riduttiva e i vari sistemi di deduzione naturale vengono presentati come strumenti per individuare il grado di evidenza dei principi usati, concentrando l'interesse sulla *dimostrabilità* entro le teorie più che sulle modalità di dimostrazione, la seconda pone al centro del discorso l'analisi *diretta* delle dimostrazioni e riconosce, nei vari sistemi di deduzione naturale, tentativi per l'analisi del concetto di prova. Il mutamento è interessante in quanto nella prima parte Kreisel aveva esplicitamente dichiarato di considerare erronea l'opinione di Gentzen, secondo la quale le regole costituiscono un'analisi del significato degli operatori logici, riconducendo in sostanza quest'analisi ad una base semantica – fosse essa insiemistica o costruttiva – che è indipendente ed anteriore alla determinazione delle regole. In questo modo l'analisi delle dimostrazioni anche se feconda sul piano tecnico risultava subordinata concettualmente ad una preliminare concezione semantica. Il riportare in primo piano, come viene fatto nel secondo articolo, il concetto di dimostrazione, significava accettarlo come elemento «naturale» dell'analisi degli operatori logici: in questo contesto l'analisi semantica, soprattutto quella in termini costruttivi o di computazione o altro, non veniva più posta come fondante, ma doveva risultare collegata puntualmente con le regole deduttive.

Un ripensamento del genere a nostro parere era in parte motivato dal sempre maggiore distacco che si era venuto a creare fra la concreta pratica matematica e ciò che la teoria della dimostrazione in senso riduttivo assume come tipico dell'esperienza matematica. Che tipo di informazioni forniscono risultati quali la dimostrazione dell'*Hauptsatz* di Gentzen, o la dimostrazione del 1958 di Gödel? Affermare che esse «giustificano» l'aritmetica o l'Analisi da un punto di vista costruttivo non sembra completamente plausibile e inoltre si ha l'impressione che la sottigliezza delle dimostrazioni «dica» qualcosa di più di questo e di diverso; informazioni del genere nel panorama della matematica come oggi viene praticata non sembrerebbero particolarmente significative: questo, si noti, anche in quei settori della matematica in cui i problemi di effettività e costruttività sono il tema stesso dell'indagine, si pensi alla dimostrazione automatica, alla problematica della *computer science*, ecc. È stato lo stesso Kreisel a porre in luce come, nel suo complesso, la prospettiva generale hilbertiana non sia sufficientemente «sofisticata» da permettere un autentico approfondimento della comprensione dei problemi posti nella pratica dalla dimostrazione automatica o dalla computazione, e non a caso, nel suo articolo *What have we learnt from Hilbert's second Problem?* (*Cosa abbiamo imparato dal secondo problema di Hilbert?*, 1976) dedicato ad un bilancio del programma hilbertiano, egli poneva l'enfasi non sulle formulazioni filosofiche generali – di carattere ideologico – ma sui risultati collaterali cui la ricerca sul programma aveva dato origine. Se questa ricerca del significato positivo del programma hilbertiano – al di là della visione sostanzialmente riduttiva della matematica che esso aveva nelle sue formulazioni generali – è una costante del modo di concepire la teoria della dimostrazione da parte di Kreisel fin dai suoi primi lavori sul significato *matematico* delle dimostrazioni di coerenza, non si può negare che dagli anni sessanta la ricerca di collegamenti più diretti con la matematica o anche con l'informatica si sia andata facendo molto più pressante nell'ambito dei ricercatori che si occupano di teoria della dimostrazione. L'enfasi – come testimoniano le ricerche di cui abbiamo riferito – è sulla struttura delle dimostrazioni, sulla possibilità di ricavare da esse informazioni – siano queste confini numerici, procedure di computo o altro – e più in generale sulla ricerca del contenuto che le dimostrazioni codificano, su come que-

sto dipenda dalla forma delle dimostrazioni e su come queste ultime si possano classificare in termini di contenuto e di efficienza.

In tempi recenti – come ancora Kreisel e per vie diverse Girard hanno mostrato – nuovi campi di applicazione si sono aperti alla teoria della dimostrazione vista come strumento per dipanare (*unwind*) specifiche dimostrazioni matematiche e ricavarne informazioni aggiuntive; ma il problema è generale e il discorso si applica anche al modo tradizionale di vedere la teoria dell'effettivo. Tutto sommato, l'interesse dei logici in questo campo è stato fino agli anni sessanta più che per l'effettivo, per il *non effettivo*, vale a dire per le dimostrazioni della non esistenza di algoritmi più che per la formulazione di procedure «semplici» per il confronto e le analisi dei programmi delle macchine. Il ritorno a un interesse più esplicitamente diretto al concetto generale di dimostrazione (per quanto riguarda la computazione si è già detto parlando delle generalizzazioni della ORT) costituisce ci sembra una presa di coscienza di questo fatto. È significativo che uno dei maggiori protagonisti delle ricerche sulla teoria generale della dimostrazione, Dag Prawitz, abbia compiuto ricerche anche nel campo della dimostrazione automatica e che logici come Scott, Erwin Engeler, ecc. abbiano elaborato e raffinato nozioni sorte nella pratica logica in modo da rendere possibile questo contatto con la pratica del calcolo automatico.

Non vogliamo peraltro sopravvalutare il ruolo di questa connessione: c'era un interesse filosofico autonomo nella ripresa di un'analisi del concetto generale di dimostrazione. Nella prospettiva di Prawitz, infatti, il ruolo di questa teoria è quello di simulare e classificare le dimostrazioni intuitive mediante le loro traduzioni formali, ed il problema centrale diveniva da una parte quello di fornire una definizione di correttezza o di verità sulla base del procedimento dimostrativo, dall'altra quello di formulare condizioni per la individuazione delle prove e quindi criteri per la loro identificazione: l'importanza filosofica di un simile programma è innegabile. In quest'ordine di idee si situano alcuni lavori di Prawitz del 1973 e del 1975 che partono dalla contrapposizione fra l'analisi semantica della correttezza degli argomenti fornita dal concetto di verità e quella che si può dare riprendendo la concezione di Gentzen delle regole di introduzione e di eliminazione come analisi del significato degli operatori logici.

Prawitz formulava una nozione di verità per le dimostrazioni che sfrutta il collegamento fra passi di riduzione e procedure di computo nel λ-calcolo di cui sopra abbiamo parlato. L'idea di fondo è che è plausibile assumere che l'equivalenza fra prove data dalla relazione di riducibilità corrisponda all'equivalenza fra le prove intuitive associate alle dimostrazioni formali. Di qui una giustificazione nuova del ruolo dei teoremi di normalizzazione visti come risultati che permettono una partizione delle dimostrazioni: due dimostrazioni sono equivalenti se hanno la stessa forma normale. In un certo senso quindi le forme normali ci permettono di individuare la complessità delle dimostrazioni intuitive e di classificarle su questa base. Queste ricerche hanno contribuito a dare una nuova formulazione al dibattito sulla natura delle leggi logiche, e in particolare al confronto fra logica classica e intuizionista. Come ad esempio è sembrato a M. Dummett, la formulazione della logica data da Prawitz rende esplicita l'idea del significato degli operatori logici come *uso* degli stessi, aprendo così la strada alla eliminazione di un approccio di tipo platonista che considera i significati come entità a sé stanti, a favore di una concezione costruttivista, quale quella dell'intuizionismo, in cui confluiscono anche elementi della visione wittgensteiniana delle regole linguistiche come regole di giochi.

2.3 *Problemi di decisione*

Come sappiamo, la teoria della dimostrazione è solo un aspetto dell'indagine generale sulle proprietà logiche delle teorie che a partire dagli anni trenta si indica col nome di metamatematica. Già nell'originale programma hilbertiano uno degli obiettivi dell'analisi delle dimostrazioni doveva essere quello di permettere di risolvere il problema della decisione per specifiche teorie. Nel capitolo precedente abbiamo ricordato come i risultati di Gödel del 1931 abbiano aperto tutta una problematica nuova in questo campo offrendo metodi per le dimostrazioni di indecidibilità. I risultati ottenuti a partire dagli anni trenta sulla decidibilità o indecidibilità delle teorie sono numerosi ed abbiamo già avuto occasione di accennarvi nel capitolo precedente, parlando della teoria dei modelli e dei rapporti tra completezza e decidibilità.

Non avrebbe senso fare un elenco dei risultati raggiunti in questo campo; ci basti dire che *grosso modo* tre sono i tipi di indagine che dagli anni cinquanta si sono andati sviluppando a questo riguardo.

Il primo riguarda la logica pura e la individuazione delle classi di riduzione per la decidibilità rispetto alla validità logica. Come il lettore ricorderà, in base al teorema di Church, non è possibile determinare in modo meccanico quando una formula è legge logica o meno, dualmente quando è soddisfacibile o meno. La domanda che ben presto si pose naturalmente fu se il problema divenisse risolubile una volta che ci si restringesse a particolari classi di formule. Centrale, ma non esclusiva, nella ricerca di queste classi è la considerazione dei prefissi, in quanto è la presenza dei quantificatori che in generale rende possibile l'indecidibilità: sul piano proposizionale, infatti, le tavole di verità ci forniscono un metodo di decisione. Già nel libro di Hilbert e Ackermann del 1928 si trovano esempi di classi di formule per cui il problema della validità è decidibile, risultati che si devono a Behmann, Bernays, Skolem e Ackermann stesso. In particolare, Behmann dimostrò nel 1922 la decidibilità del problema della validità per formule contenenti solo predicati monadici e lo stesso vale per formule del tipo Π_2 che non contengono simboli funzionali.

I risultati più forti in questa direzione si devono a Gödel ed a Herbrand il quale ultimo utilizzò in questo contesto i suoi metodi particolari di analisi delle dimostrazioni. Nel 1932 Gödel dimostrò la risolubilità del problema della validità per formule con prefisso del tipo

$$\forall x_1,\ldots,x_n \, \exists y \exists z \forall y_1,\ldots,\forall y_l$$

dopo che Ackermann l'aveva dimostrato nel 1928 per formule con prefisso

$$\forall \ldots \, \forall \exists \forall \ldots \, \forall$$

ossia con un solo esistenziale. Si otteneva così un confine piuttosto netto tra decidibilità e indecidibilità: due quantificatori esistenziali contigui non sono sufficienti per darci l'indecidibilità mentre tre – si può dimostrarlo – lo sono.

Da parte sua Herbrand aveva provato nel 1930 che il problema della decisione era risolubile considerando limitazioni che non coinvolgono solo il prefisso e che riguardano le formule il cui

prefisso ha quantificatori universali arbitrari ma la cui matrice contiene solo negazione e disgiunzione. I metodi usati in queste dimostrazioni consistono essenzialmente nella riduzione del problema della validità rispetto a tutte le interpretazioni, a quello rispetto a una classe ristretta per la quale si possa stabilire un confine superiore finito alla cardinalità del dominio. Un'esposizione sistematica di questi metodi venne data nel 1954 da Ackermann nel suo libro *Solvable Cases of the Decision Problem* (*Casi risolubili del problema della decisione*). Più complesso era il problema di localizzare, per così dire, dal versante della indecidibilità, le classi di formule più semplici per cui il problema della validità non è risolubile.

Si giungeva così a considerare le cosiddette *classi di riduzione*, cioè quelle classi di formule a cui è possibile ricondurre il problema della decisione per ogni formula. È chiaro che le classi di riduzione sono indecidibili ed il problema interessante è quello di determinare quali sono quelle di forma più semplice. I risultati più interessanti al riguardo furono ottenuti nel 1950 da Janos Suranyi che mostrò che le formule con prefisso $\exists\exists\exists\forall$ hanno problema della validità indecidibile, risultato perfezionato poi con metodi diversi da Kahr, Moore e Wang, che dimostrarono la indecidibilità delle formule con prefisso del tipo $\exists\forall\exists$, riconducendo il problema della decisione di queste formule alla soluzione del problema del domino.

Collegando questi risultati con quelli precedenti si otteneva una classificazione esaustiva per quanto riguarda la decidibilità delle classi di formule determinate dal tipo di prefisso così che il problema della validità per formule di un dato prefisso è indecidibile *se e solo se* il prefisso o presenta due quantificatori esistenziali separati da un universale ($\exists\forall\exists$) o almeno tre quantificatori esistenziali che precedono un quantificatore universale ($\exists\exists\exists\forall$).

Il metodo utilizzato da Gödel si basava su un criterio trovato in modo indipendente anche da Kalmar e da Schütte, e che permetteva di affrontare una questione più generale di quella della decidibilità, quella della controllabilità finita. Si dice che una classe di formule è *finitamente controllabile* se ogni formula soddisfacibile della classe ha un modello finito. È chiaro che il problema della soddisfacibilità è decidibile per una data classe, se essa è finitamente controllabile; il viceversa non è sempre vero. Gödel, nel corso del-

la sua dimostrazione, provava la controllabilità finita della classe delle formule con prefisso $\forall\forall\exists \ldots \exists$ da cui si ottiene la controllabilità finita della classe delle formule con prefisso $\exists \ldots \forall\forall\exists \ldots \exists$.

La conclusione generale è che nel caso di classi di formule determinate dalla forma del prefisso, controllabilità finita equivale a decidibilità del problema della soddisfacibilità, cosicché la considerazione delle interpretazioni finite è un metodo adeguato per il problema in questione. Nel suo lavoro Gödel affermava che il metodo si poteva estendere anche nel caso di formule che contenevano il simbolo $=$, cosa che è stata dimostrata non essere vera, in tempi recenti, da Warren Goldfarb, e che ha dato origine a nuove ricerche sull'argomento.

Se la considerazione delle classi di riduzione riguarda la logica pura, dagli anni trenta in poi – come già ricordato – si sono succedute numerose ricerche su decidibilità o indecidibilità di teorie matematiche *specifiche* e molto attiva è stata la ricerca di metodi generali per affrontare sia problemi di decidibilità sia problemi di indecidibilità. Come si ricorderà, una delle strade maestre per provare la decidibilità è quella che passa attraverso la assiomatizzabilità e la completezza e forse il risultato più interessante ottenuto negli anni sessanta è appunto la decidibilità della teoria dei numeri p-adici raggiunta da James Ax e S. Kochen nel 1965, nel contesto di un'indagine generale sui campi valutati, il cui risultato più significativo dal punto di vista matematico è la dimostrazione definitiva, via completezza, della verità di una congettura di Artin. L'interesse del risultato sta nel fatto che costituisce il primo esempio di teorema puramente algebrico dimostrato *direttamente* con i metodi della logica matematica, nel senso che tutte le dimostrazioni che di esso sono state date sono modificazioni della dimostrazione data in termini di teoria dei modelli. Gli stessi Ax e Kochen dimostrarono nel 1968 la decidibilità della teoria dei campi finiti aprendo la strada a interessanti ricerche in questo settore.

La tecnica utilizzata per la decidibilità sfruttava la completezza ma è interessante osservare che in molti casi si può procedere in maniera diretta dalla eliminazione dei quantificatori (E.Q.) alla decidibilità, una volta che si provi la risoluzione del problema della decisione all'interno della teoria in esame relativamente alla classe delle formule prive di quantificatori, in quanto appunto

questa classe costituisce ora una classe di riduzione rispetto alla teoria. Questo collegamento, su cui ci siamo già fermati nel capitolo precedente, pone in primo piano il problema della E.Q. di cui negli anni sessanta, sulla scorta di precedenti lavori di A. Robinson, Shoenfield e A. MacIntyre, è stato indagato il significato algebrico, non solo per le varie teorie dei campi, ma anche per quelle dei gruppi e dei moduli.

Dal punto di vista dei problemi della decisione in senso stretto, i risultati più interessanti riguardano però la decidibilità di teorie di ordine superiore al primo: così ad esempio nel 1960 J. Richard Büchi dimostrava che l'aritmetica del secondo ordine con la sola operazione di successore è decidibile. Questo risultato è molto interessante in quanto questa teoria permette di esaminare in modo duttile gli automi finiti (le macchine sequenziali). Nella stessa direzione si muovevano le ricerche di John Doner che ha esteso il risultato di Büchi e il cosiddetto teorema dell'albero di M. Rabin al caso di un'aritmetica al secondo ordine con un numero finito arbitrario di operazioni di successore. Una teoria come quella studiata da Büchi si può vedere come un'assiomatizzazione degli insiemi bene ordinati con una sola operazione di successore; introducendo più operazioni di successore è possibile interpretare nel sistema risultante la teoria di tipi d'ordine più generali e ottenerne così, in vista del risultato di Doner, la decidibilità. I risultati sono poi stati estesi da Hans Läuchli alla teoria degli ordini lineari formulata in un linguaggio del secondo ordine debole.

Negli anni cinquanta e sessanta si ottennero anche interessanti risultati sui metodi *generali* utilizzabili per dimostrare la decidibilità e indecidibilità che si affiancano a quelli di cui abbiamo già parlato. Tra questi possiamo ricordare il metodo di Alan Cobham presentato nel 1966, che permette di ricondurre l'indecidibilità di una teoria $\mathfrak{T}$ alla sua compatibilità con un frammento particolarmente debole dell'aritmetica, la teoria $\mathfrak{P}_0$ e che è stata successivamente indagato da Robert Vaught. Accanto a questo va ricordato il perfezionamento della tecnica delle interpretazioni di Tarski, presentato da Rabin nel 1965, che collega l'indecidibilità di una teoria $\mathfrak{T}$ alla possibilità di rappresentare sintatticamente al suo interno, ed in modo effettivo, ogni modello di una teoria indecidibile $\mathfrak{T}'$.

Sul versante della decidibilità, di estremo interesse sono invece

le ricerche sui prodotti di teorie iniziate da Mostowski già nei primi anni cinquanta e che confluirono nell'articolo fondamentale di Feferman e Vaught del 1959 sulle proprietà elementari dei prodotti di strutture algebriche, dal titolo *The first order properties products of algebraic systems* (*Le proprietà del primo ordine di sistemi algebrici*). In questo lavoro gli autori offrivano un contesto unitario in cui presentare operazioni disparate quali i prodotti diretti, i prodotti ridotti, le somme disgiunte, ecc. e mostrarono come la verità di una formula elementare su uno qualsiasi dei risultati di queste operazioni fosse riconducibile alla verità nelle strutture di partenza. Il collegamento col problema della decisione nasceva dal fatto che, una volta che si sappia che tutti i modelli di una data teoria si ottengono con un dato tipo di prodotto dai modelli di un'altra, è possibile risolvere il problema della decisione per la prima teoria rimandandolo alla seconda. Così ad esempio dal risultato di Presburger sulla decidibilità dell'aritmetica con la sola operazione di addizione, possiamo ottenere il risultato di Skolem che stabilisce la decidibilità della teoria degli interi visti come semigruppo moltiplicativo.

Più che soffermarci sulle diverse applicazioni che questi metodi hanno permesso e sulle loro ramificazioni, vogliamo però parlare ora di alcuni tipi di risultati di indecidibilità che riguardano teorie particolari e che hanno aperto la strada a nuovi tipi di collegamento fra logica e matematica.

Il primo è il teorema di S.A. Cook, dimostrato nel 1971, che riconduce al problema della soddisfacibilità per la logica proposizionale il problema tutt'oggi aperto, e che ha un ruolo centrale nella teoria della complessità, $P \stackrel{?}{=} NP$ che consiste nel chiedersi se coincida la classe dei problemi risolubili in tempo polinominale in modo non deterministico (NP) con quella dei problemi risolubili sempre in tempo polinomiale con algoritmi deterministici. Il teorema di Cook afferma che vale $P = NP$ nel caso sia risolubile in tempo polinomiale il problema della decisione per il calcolo proposizionale. Dopo il lavoro di Cook le ricerche in questa direzione si sono moltiplicate e numerosissimi sono stati i problemi combinatori di decisione la cui risolubilità in tempo polinomiale è stata provata equivalente a $P \stackrel{?}{=} NP$. Questi tipi di indagine hanno segnato l'inizio di un approccio più realistico ai problemi della decisione rispetto a quello tradizionale del logico, in

cui la semplice domanda: la teoria $\mathfrak{T}$ è decidibile? è sostituita dall'interrogativo più sottile: $\mathfrak{T}$ è decidibile in tempo polinomiale? O in generale, che tipo di complessità ha il problema?

Di portata diversa sono gli altri due risultati concernenti la decidibilità di cui ora vogliamo parlare, sia perché rispondono ad interrogativi antichi, sia perché hanno aperto la strada a campi di ricerca di particolare interesse. Il primo è il risultato ottenuto da Yuri V. Matijasevič nel 1970, relativamente al decimo problema di Hilbert. Esso stabilisce la non esistenza di un algoritmo in grado di decidere la risolubilità nel dominio degli interi delle equazioni diofantee. La sostanza del risultato di Matijasevič sta nella dimostrazione del fatto che i predicati *diofantei* di numeri naturali $\mathscr{A}(y_1, ..., y_n)$ coincidono con i predicati ricorsivamente enumerabili. I predicati diofantei si ottengono in questo modo. Sia $P(y_1, ..., y_n, x_1, ..., x_k)$ un polinomio con coefficienti interi $y_1, ..., y_n$. Possiamo considerare la formula

$$\exists x_1, ..., x_k \; P(y_1, ..., y_n, x_1, ..., x_k) = 0$$

come una definizione dell'insieme delle n-uple $< y_1, ..., y_n >$ di numeri naturali, per cui l'equazione ha una soluzione *nell'insieme dei numeri naturali*, nelle incognite $x_1, ... x_k$. Quindi, come esistono predicati ricorsivamente enumerabili che non sono ricorsivi, esisteranno predicati «diofantei» che non sono ricorsivi. Questo significa appunto che – se si identifica computabile con ricorsivo – non esiste l'algoritmo richiesto da Hilbert. Oltre a risolvere un problema vecchio di mezzo secolo, il risultato di Matijasevič gettava nuova luce sullo studio dei modelli dell'aritmetica di Peano, studio in cui teoria della ricorsività e semantica si legano strettamente. Alla base di questo sta la semplice osservazione fatta da Haim Gaifman nel 1970 che sulla base del teorema di Matijasevič possiamo equiparare le relazioni ricorsive con quelle *assolute* rispetto alle estensioni elementari della struttura dei naturali. In altre parole, una formula $\mathscr{A}$ interpretata nella struttura dei naturali definisce una relazione ricorsiva se e solo se prese due strutture $\mathfrak{A}$ e $\mathfrak{B}$ con $\mathfrak{A} \subseteq \mathfrak{B}$ entrambe estensioni elementari di N, per ogni $a_1, ..., a_n \in \mathfrak{A}$ si ha che

in $\mathfrak{A}$ è vera $\mathscr{A}(a_1, ..., a_n)$ se e solo se in $\mathfrak{B}$ è vera $\mathscr{A}(a_1, ..., a_n)$.

La caratterizzazione così data delle funzioni e quindi delle relazioni ricorsive risulta più sottile di quella in termini di definibilità invariante di cui abbiamo già parlato, in quanto riguarda la verità in estensioni elementari *qualsiasi* di N di formule su individui che possono *non* appartenere ad N.

Questo collegamento tra equazioni diofantee ed estensioni elementari di N non era né casuale né inaspettato. Già nel 1962 M. Rabin aveva provato che per ogni estensione elementare di N esiste un'equazione diofantea con coefficienti nel modello, che non ha soluzione nel modello dato, ma in una sua estensione opportuna. Il fatto è interessante in quanto questo non può succedere per equazioni diofantee con coefficienti che siano numeri naturali. Di questo Rabin si serviva per dare una nuova dimostrazione del teorema di Ryll-Nardzewski sulla non assiomatizzabilità finita dell'aritmetica di Peano, dimostrazione che darà origine a interessanti sviluppi nello studio dei modelli non standard dell'aritmetica. Ma di questo avremo occasione di parlare più avanti.

Vogliamo ora invece considerare un risultato che riguarda sempre l'aritmetica e che ha segnato una vera svolta nella considerazione della indecidibilità. Come sappiamo, il primo teorema di Gödel dimostra l'esistenza di almeno un enunciato vero nella struttura dei naturali ma non dimostrabile dagli assiomi di Peano. Il contenuto di questo enunciato – se dal punto di vista metamatematico è significativo – non lo è dal punto di vista matematico, in quanto non stabilisce alcuna proprietà «interessante» sui numeri naturali. È questa una delle non ultime ragioni che hanno portato spesso i matematici a sottovalutare il significato del teorema stesso e la domanda naturale era se fosse possibile trovare enunciati veri non dimostrabili in $\mathfrak{P}$ matematicamente significativi.

Nel 1976 J. Paris e L. Harrington fornirono per la prima volta un esempio di enunciato di questo tipo. Si trattava di una modificazione della versione finita di quel teorema di Ramsey di cui abbiamo parlato a suo tempo, sottolineandone il ruolo centrale nella combinatorica, e che afferma che una volta che si definisca *relativamente grande* un insieme finito di naturali H quando la sua cardinalità è maggiore o uguale al minimo dei numeri in esso contenuti, data una partizione $P : [M]^e \rightarrow r$ dove e, r e M sono naturali, allora esiste un sottoinsieme H di numeri minori di M relativamente

grande omogeneo per P e di cardinalità almeno k per k prefissato.[4] Essenziale è la clausola «relativamente grande»; senza di essa l'enunciato che afferma che quanto sopra si verifica per ogni e, r, k e M coincide con la versione finita del teorema di Ramsey, e come tale si può provare che è dimostrabile in $\mathfrak{P}$. Quello che in $\mathfrak{P}$ non è dimostrabile è la nostra versione, che indicheremo con E.

Originariamente Paris e Harrington dettero una dimostrazione sintattica modellata sull'esempio del teorema di Gödel e solo più tardi ne dettero una versione in termini di teoria dei modelli sfruttando alcuni fatti fondamentali sui modelli non standard dell'aritmetica. La strategia seguita dagli autori è semplice: per prima cosa essi introducono una teoria $\mathfrak{T}$, nel linguaggio dell'aritmetica, arricchita con infinite nuove costanti individuali; dimostrano in $\mathfrak{P}$ che Coer($\mathfrak{T}$) implica Coer($\mathfrak{P}$) e provano quindi che d'altra parte, sempre in $\mathfrak{P}$, E implica Coer($\mathfrak{T}$). Il secondo teorema di Gödel ci dà immediatamente la non dimostrabilità di E.

La portata del risultato appena citato emerge con particolare rilievo una volta che si consideri che il principio combinatorio E non è una proprietà scelta *ad hoc*, ma ha un grosso significato nello studio di quei processi combinatori finiti che sono codificabili all'interno dell'aritmetica, cosicché viene naturale la domanda di quali siano i teoremi matematici conseguenza del principio, la cui non dimostrabilità scende dalla non dimostrabilità di E e all'inverso di quali altri risultati aritmetici esso sia conseguenza. Lungo questa via si giunge a tutta una serie di interrogativi che costituiscono il nucleo di quella che, a partire dal 1974, Harvey Friedman ha battezzato la *reverse mathematics* e che – pur se le reazioni sono state discordi – costituisce un indirizzo di ricerca di particolare interesse. L'idea di Friedman è quella di indagare sullo sfondo di una teoria di riferimento – nei casi che a noi interessano un particolare sottosistema dell'aritmetica del secondo ordine – quali sono i principi sufficienti, ma anche *necessari* per dimostrare risultati di particolare significato matematico. Con le parole che Friedman usò nella sua comunicazione al congresso dei matematici del 1974: «Quali sono gli assiomi propri da usare per dimostrare particolari

[4] Al solito con la notazione $P:[M]^e \rightarrow r$ indichiamo una partizione in r classi disgiunte della famiglia degli insiemi di numeri naturali minori di M aventi ciascuno cardinalità uguale ad e.

teoremi o insiemi di teoremi nella matematica? Quali sono i sistemi formali che ci permettono di isolare i principi essenziali per dimostrarli?» Friedman è stato il primo a mostrare come sia possibile studiare all'interno di sottosistemi dell'aritmetica del secondo ordine ampi settori della matematica non insiemistica.

Già Hilbert e Bernays avevano sottolineato come la maggior parte dei teoremi della matematica predicativista elementare fosse dimostrabile in $\mathfrak{P}^2$ e Weyl prima e Kreisel nel 1959 avevano analizzato entro che limiti fossero adeguati sottosistemi di tipo predicativo, ad esempio limitandosi a postulare l'assioma di comprensione per predicati iperaritmetici. Lungo questa linea Friedman ha potuto ad esempio dimostrare che considerando il sottosistema RCA, particolarmente debole, in cui si postula l'assioma di comprensione solo per i predicati Δ_0 (cioè i predicati ricorsivi) si ha che la versione debole del teorema di König (ogni albero infinito di successioni di 0 e 1 ha almeno un ramo infinito) è condizione non solo sufficiente ma anche necessaria per dimostrare la compattezza per la logica proposizionale, la completezza del calcolo dei predicati, il lemma di Lindenbaum (tutto questo ovviamente per linguaggi numerabili) e il teorema di Heine-Borel in forma sequenziale. Da allora le ricerche in questa direzione si sono moltiplicate e hanno permesso di vedere come, scegliendo opportunamente il sistema di riferimento, si possa indagare in casi molto significativi quali principi sono necessari e quali no. Anche in questo caso – come Friedman ha sottolineato – l'aspetto interessante è che la forza dei principi analizzati è misurata non in termini astratti, ma in riferimento a problemi che si pongono nella pratica matematica ordinaria (che non assume cioè come oggetti di studio direttamente insiemi o strutture arbitrarie).

Anch'essi antichi – la loro origine risale ai primi decenni del secolo e alle ricerche del topologo Max Dehn – sono altri problemi che si presentano nell'algebra e nella topologia. Ci riferiamo ai cosiddetti *problemi della parola*, che in generale hanno la forma seguente: dato un alfabeto finito A e delle regole di trasformazione per le parole costruite su A, trovare un algoritmo in grado di determinare quando per due qualsiasi parole t e t' è dimostrabile l'eguaglianza $t = t'$. Come si vede si tratta di un problema di decisione per un particolare sistema formale i cui teoremi sono le formule del tipo $t = t'$, dove t e t' sono parole sull'alfabeto. Se la teo-

ria generale di sistemi di questo tipo era già stata sviluppata da Axel Thue già nel 1914, problemi del genere divennero vitali per la matematica convenzionale, quando si considerò che semigruppi e gruppi si possono presentare come sistemi di questo tipo dove l'operazione di concatenazione fra parole corrisponde all'operazione di gruppo o semigruppo e come equazioni di base (assiomi) si prendono le equazioni definitorie del gruppo o semigruppo. Come è noto infatti ogni gruppo si può presentare come quoziente di un gruppo libero i cui elementi sono appunto parole e la cui operazione è la concatenazione; le equazioni definitorie sono le equazioni che «generano» la congruenza rispetto alla quale si fa il quoziente utilizzando le regole sull'identità.

All'origine del problema della parola – come dei collegati problemi del coniugio e dell'isomorfismo – stava il rapporto tra spazi topologici puntati e gruppi fondamentali. In questi contesti i gruppi fondamentali si offrono naturalmente in termini di presentazioni e la questione centrale è quella di estrarre dalla presentazione informazioni sulla struttura del gruppo e quindi degli spazi associati. Nel 1955 P.S. Novikov era riuscito a dimostrare che esistevano gruppi finitamente presentati in cui il problema della parola era insolubile, così che non c'era un metodo meccanico per decidere quando due parole designavano lo stesso elemento. La dimostrazione di Novikov era estremamente complessa e dopo diversi tentativi (uno dei quali dovuto a Turing) nel 1959 W.W. Boone dava una nuova dimostrazione dello stesso risultato. Al fondo delle dimostrazioni tanto di Novikov che di Boone stava l'idea di associare ad ogni gruppo finitamente presentato un'opportuna macchina di Turing cosicché il problema della parola per il gruppo si trasformasse nel problema della fermata per quella macchina. In questa maniera non si sfruttava gran che della teoria dei gruppi e si scaricava tutto sullo studio delle macchine di Turing o strutture combinatorie equivalenti. Era questa la linea che era stata seguita precedentemente da Thue, Post e Markov per l'analogo problema relativo ai semigruppi, la cui indecidibilità Post e Markov avevano dimostrato alla fine degli anni quaranta.

Un mutamento decisivo di prospettiva avvenne nel 1963 quando J.L. Britton mostrò come le complicate costruzioni combinatorie si potevano rimpiazzare utilizzando teoremi algebrici sui gruppi riguardanti i prodotti amalgamati. Sempre con questi stessi

strumenti Graham Higman – cui si devono alcuni risultati fondamentali sui prodotti amalgamati, ottenuti in collaborazione con Hanna e Bernhard Neumann – dimostrò nel 1961 un teorema che non solo aveva tra gli immediati corollari quello di indecidibilità di Boone e Novikov, ma mostrava come un problema squisitamente algebrico sui gruppi finitamente presentati ammettesse una soluzione in termini di teoria della ricorsività. Il teorema di Higman afferma infatti che un gruppo si immerge in un gruppo finitamente presentato se e solo se ha una presentazione ricorsiva, in cui cioè generatori e relazioni definitorie costituiscano insiemi ricorsivi. In questo modo la proprietà algebrica di essere immergibile in un gruppo finitamente presentato si rivelava *equivalente* alla proprietà logica di avere presentazione ricorsiva. Sempre fondandosi sull'analogia fra presentazioni di gruppi e teorie, nel 1973 Boone e Higman potevano caratterizzare i gruppi finitamente generati con problema della parola risolubile come quelli immergibili in sottogruppi semplici di gruppi finitamente presentati, dove l'analogia chiave è tra teorie complete e gruppi semplici.

L'indecidibilità del problema della parola ha controparti topologiche, come è facile immaginare e come già negli anni trenta Church aveva prefigurato. Così ad esempio nel 1958 Markov poteva dimostrare l'insolubilità del problema della omeomorfia per varietà di dimensioni maggiori o uguali a quattro (si ricordi che è un risultato classico che risale a Riemann, la soluzione positiva del problema nel caso di due dimensioni). Collegato al problema dell'omeomorfia – sfruttando il passaggio ai gruppi fondamentali – è quello dell'isomorfismo: quando, date due presentazioni, i gruppi così presentati sono isomorfi? Si può provare, come corollario immediato del teorema di Boone-Novikov, che in generale il problema non è risolubile per gruppi finitamente presentati e nel 1955 Adjan e nel 1958 Rabin stabilivano un risultato generale secondo il quale non esiste un metodo per decidere quando un gruppo con una data presentazione appartiene ad una determinata classe se questa è chiusa rispetto all'isomorfismo, non banale e chiusa rispetto ai sottogruppi. Da questo risultato discende non solo quello precedente, ma pure che non c'è metodo per decidere dalla presentazione se un gruppo è finito, libero, commutativo, ecc.

Nel corso del tempo questi metodi e questi risultati sono stati

estesi in molte direzioni, sia nel senso che si sono affrontate altre questioni riguardanti le presentazioni (primo fra tutti il classico *problema di Burnside*, che risale al 1902, di sapere se ogni gruppo con un numero finito di generatori che soddisfa la relazione $x^n = 1$ è finito) sia nel senso che si è cercato di estendere quanto fatto per i gruppi alle presentazioni di altre strutture, ad esempio gli anelli e i corpi. Un tipo di sviluppo diverso è emerso invece all'inizio degli anni settanta e riguarda il collegamento tra gruppi algebricamente chiusi e problema della parola. Un gruppo è algebricamente chiuso se ogni insieme di equazioni su di esso definito ha una soluzione già nel gruppo stesso se ha soluzioni in una sua estensione, cioè se è «compatibile». Il problema della risolubilità di equazioni su gruppi è estremamente complesso e fu affrontato già negli anni trenta da B. Neumann; l'idea era che i gruppi algebricamente chiusi avrebbero svolto nella teoria dei gruppi un ruolo analogo a quello dei campi algebricamente chiusi nella teoria dei campi, in quanto è facile provare che ogni gruppo si immerge in un gruppo algebricamente chiuso.

Nel 1973 Neumann poneva il problema fondamentale di trovare criteri per determinare quando due gruppi numerabili algebricamente chiusi sono isomorfi e congetturò che il problema fosse insolubile. Nel corso delle indagini sui gruppi algebricamente chiusi Neumann riuscì a stabilire che ogni gruppo finitamente generato che ammetta presentazione con problema della parola risolubile, è immergibile in tutti i gruppi algebricamente chiusi. Il fatto interessante dimostrato da A. MacIntyre è che vale anche l'inverso, e che se un gruppo finitamente presentato ha problema della parola non risolubile, allora esiste un gruppo algebricamente chiuso in cui non si può immergere. Si otteneva così una caratterizzazione *puramente algebrica* dei gruppi finitamente presentati con problema della parola risolubile e MacIntyre poteva provare, ritornando al problema originario di Neumann, che due gruppi algebricamente chiusi numerabili sono isomorfi se e solo se hanno sottogruppi finitamente generati che realizzano gli stessi tipi di isomorfismi. Il problema di Neumann era essenzialmente un problema di isomorfismi, e come tale si può indebolire, ad esempio chiedendosi se due gruppi algebricamente chiusi sono necessariamente elementarmente equivalenti. È questo il problema di E. Bers formulato negli anni settanta e che a tutt'oggi non ha avuto

risposta, come pure un altro classico problema analogo posto da Tarski sui gruppi liberi: tutti i gruppi liberi con almeno due generatori sono tra loro elementarmente equivalenti?

Tutti questi interrogativi si possono formulare in termini di completezza di teorie e di associati problemi di decisione e non si pongono solo sui gruppi ma più in generale su strutture non commutative, come i corpi, le algebre di Lie. ecc.; al di là degli specifici risultati, quello che emerge è uno stretto intergioco tra teoria dei modelli, algebra e teoria della ricorsività. Ed è appunto alla teoria dei modelli che vogliamo ora rivolgerci.

2.4 *Model-theoretic algebra e linguaggi infinitari*

I problemi di completezza e decidibilità, pur importanti, costituiscono solo un aspetto della problematica affrontata a partire dagli anni sessanta dalla teoria dei modelli e le direzioni più significative di ricerca riguardano altre questioni: lo sviluppo della teoria di Morley sulla categoricità in potenza, l'approfondimento del concetto di model-completezza per mezzo della nozione di *forcing* generalizzato introdotta da Robinson e più in generale la *model-theoretic algebra*, la teoria dei modelli per linguaggi che estendono quelli elementari: linguaggi infinitari, con quantificatori generalizzati, ecc.

Abbiamo già avuto occasione di parlare dei risultati di Morley sulla categoricità nel capitolo VI, 2. L'importanza di questi risultati sta nella luce che essi hanno gettato sulla categoricità in potenze più che numerabili. Per la $\aleph_0$-categoricità, infatti, la situazione era più semplice ed esisteva al riguardo un fondamentale risultato di Ryll-Nardzewski del 1959 che offriva una caratterizzazione delle teorie $\aleph_0$-categoriche, e le identificava con quelle che per ogni numero naturale n posseggono un numero finito di tipi nelle variabili $x_1, \ldots, x_n$. Agli inizi degli anni sessanta risalgono altri importanti risultati al riguardo ottenuti soprattutto da R. Vaught.

Le tecniche di Morley aprivano orizzonti completamente nuovi, sfruttando e rafforzando precedenti risultati tipo quelli sui *due cardinali* (quando una struttura $\mathfrak{A}$ per un linguaggio con una certa costante predicativa P ha un'estensione elementare $\mathfrak{B}$ di potenza fis-

sata e in cui l'insieme degli oggetti che godono di P soddisfa prefissate condizioni di cardinalità?) e quelli sugli indiscernibili dimostrati da Ehrenfeucht e Mostowski nel 1956. Questi ultimi risultati riguardavano la possibilità di costruire modelli con un insieme infinito di automorfismi, iniettando in essi insiemi linearmente ordinati che inducono gli automorfismi in questione. Così ad esempio il teorema principale dimostrato da Ehrenfeucht e Mostowski afferma che se $\mathfrak{T}$ è una teoria in un linguaggio numerabile con almeno un modello infinito e X un insieme linearmente ordinato e infinito, esiste un modello $\mathfrak{A}$ di $\mathfrak{T}$ che contiene X e tale che ogni automorfismo di X induce un automorfismo su $\mathfrak{A}$. Risultati di questo tipo si sono moltiplicati dopo quelli citati e coinvolgono tutti in modo essenziale teoremi di partizione come quello di Ramsey.

Tutte queste tecniche venivano sfruttate da Morley all'interno di un'indagine sistematica sulle estensioni delle strutture e sui tipi di isomorfismo. L'idea base, grosso modo, è la stessa che ritroviamo nella teoria della estensione dei campi. Mediante un opportuno arricchimento del linguaggio è possibile, data una teoria $\mathfrak{T}'$, ricondurre lo studio dei suoi modelli a quello dei modelli di una teoria $\mathfrak{T}$ *subcompleta* tale cioè che per ogni sottostruttura $\mathfrak{A}$ di un modello di $\mathfrak{T}$, $\mathfrak{T} \cup D(\mathfrak{A})$, dove $D(\mathfrak{A})$ è il diagramma di $\mathfrak{A}$, è una teoria completa. Per studiare i tipi di isomorfismo dei modelli di $\mathfrak{T}'$ è sufficiente quindi concentrare l'attenzione sui modelli di $\mathfrak{T}$. Sia ora $\mathfrak{A}$ un sottomodello di un modello di $\mathfrak{T}$. Possiamo immaginare le estensioni di $\mathfrak{A}$ che sono modello di $\mathfrak{T}$ come ottenute mediante aggiunzione di un singolo elemento per volta. (È chiara l'analogia con la teoria dell'estensione dei campi). Chiamiamo estensione *semplice* di $\mathfrak{A}$ un'estensione $\mathfrak{A}[b]$ se è generata da $\mathfrak{A} \cup \{b\}$. Ancora in analogia con la teoria dei campi possiamo dire che due estensioni semplici $\mathfrak{B}$ e $\mathfrak{C}$ di $\mathfrak{A}$ sono *isomorfe su* $\mathfrak{A}$ se esiste un isomorfismo tra di esse che lascia invariati gli elementi di $\mathfrak{A}$. Lo scopo della teoria di Morley è quello di classificare i tipi di isomorfismo delle estensioni semplici di $\mathfrak{A}$ così che risulti possibile affrontare il problema dell'isomorfismo tra i modelli della teoria data. L'idea è di classificare questi tipi di isomorfismo estendendo la distinzione che esiste nella teoria dei campi fra elementi algebrici e trascendenti assegnando un grado di «trascendenza» agli elementi delle estensioni di $\mathfrak{A}$.

Si tratta di considerare i tipi che questi elementi realizzano. Sostanzialmente, un *tipo* su $\mathfrak{A}$ (o compatibile con $\mathfrak{A}$) è un insieme non

contraddittorio e massimale di formule con una e una sola variabile libera formulate nel linguaggio $L_{\mathfrak{A}}$ che possiede nomi per gli elementi di $\mathfrak{A}$ e tale che ogni suo sottoinsieme finito è soddisfatto da almeno un elemento di $\mathfrak{A}$ o di una estensione di $\mathfrak{A}$.[5] Più in generale se consideriamo *teorie* al posto di singoli modelli, un n-tipo è un insieme di formule con libere le sole variabili $x_1, \ldots, x_n$ massimale e coerente e si dice che esso è compatibile con la teoria $\mathfrak{T}$ se per ogni sua parte finita X esiste un modello $\mathfrak{A}$ di $\mathfrak{T}$ in cui una n-upla di elementi $<a_1, \ldots, a_n>$ *realizza*, ossia soddisfa, tutte le formule di X. È facile verificare, applicando il teorema di compattezza, che per ogni teoria $\mathfrak{T}$, se un tipo è con essa compatibile, esiste un modello di $\mathfrak{T}$ che lo realizza. È in questo modo che si costruiscono i modelli saturi, speciali, ecc., sfruttando la possibilità di costruire catene di estensioni elementari, a partire da una struttura data, che realizzano sempre più tipi. Duale a quello di realizzazione è il concetto di *omissione* di un tipo che presenta problemi notevolmente complessi: si dice che una struttura $\mathfrak{A}$ *omette* un n-tipo quando non esiste in $\mathfrak{A}$ nessuna n-upla che lo realizzi. Così ad esempio, un campo ordinato $\mathfrak{A}$ è archimedeo, quando omette il tipo costituito dalle formule

$$x > 1, x > 2, \text{ecc.}$$

La domanda naturale è: quando una teoria $\mathfrak{T}$ ha un modello che omette un dato tipo? L. Henkin nel 1954 e S. Orey nel 1956 avevano dimostrato che questo si verifica nel caso $\mathfrak{T}$ sia in un linguaggio numerabile e ometta *localmente* il tipo, vale a dire che non c'è nessuna formula, nelle variabili libere $x_1, \ldots, x_n$ coerente con $\mathfrak{T}$ che implica *tutte* le formule del tipo. Il teorema, noto come teorema di *omissione dei tipi*, svolge un grosso ruolo nello studio dei modelli delle teorie complete, e quindi della categoricità, e non ammette generalizzazioni a linguaggi più che numerabili. Lavori successivi, in particolare di R. Vaught, hanno indagato sistematicamente il problema dell'omissione dei tipi in connessione con lo studio dei modelli *primi* ed *atomici*, modelli che in un certo senso sono duali a quelli saturi in quanto, per così dire, realizzano il *minor* numero possibile di tipi.

[5] Non si confondano questi *tipi*, che sono particolari classi di *formule* in un linguaggio, con i *tipi di isomorfismo*, che sono classi di *strutture* fra loro isomorfe. Per evitare confusioni scriveremo da ora in poi «i-tipo» per «tipo di isomorfismo».

Ma torniamo a Morley e al modo di classificare le estensioni semplici utilizzando i tipi sulla struttura $\mathfrak{A}$.

In base a un semplice argomento di compattezza si può provare che se P è un tipo compatibile con $\mathfrak{A}$ esisterà un'estensione semplice $\mathfrak{A}[b]$ di $\mathfrak{A}$ tale che *b realizza P in ogni estensione di* $\mathfrak{A}$ *modello di* $\mathfrak{T}$. Questo discende dal fatto che $\mathfrak{T}$ è subcompleta. È facile provare che se tanto b che c realizzano P esisterà un unico isomorfismo $\mathfrak{A}[b] \rightarrow \mathfrak{A}[c]$ su $\mathfrak{A}$ per cui b viene mandato in c. All'inverso è ovvio che se $\mathfrak{A}[b]$ e $\mathfrak{A}[c]$ sono isomorfi su $\mathfrak{A}$ rispetto a un isomorfismo che manda b in c, b e c realizzeranno gli stessi tipi. Ne concludiamo quindi che nel caso $\mathfrak{T}$ sia subcompleta gli i-tipi delle estensioni semplici di $\mathfrak{A}$ sono univocamente individuati dai tipi compatibili con $\mathfrak{A}$. Il passo successivo sta nell'indurre una struttura sufficientemente ricca sulla famiglia $S(\mathfrak{A})$ dei tipi compatibili con $\mathfrak{A}$. Si tratta di una struttura topologica: è infatti possibile definire una topologia su $S(\mathfrak{A})$ prendendo come base gli insiemi

$$U_{\mathscr{A}(x)} = \{P \mid \mathscr{A}(x) \in P\}$$

dove $\mathscr{A}(x)$ è una formula di $L_{\mathfrak{A}}$ con libera la sola x. Lo spazio risultante è compatto e di Hausdorff. Ridotta la classificazione degli i-tipi a quella dei tipi compatibili con $\mathfrak{A}$, il problema è quello di classificare i punti dello spazio $S(\mathfrak{A})$. Senza entrare in particolari, diremo che Morley riesce ad assegnare ad ognuno di questi punti un *grado* e un *rango*: i ranghi sono ordinali e i gradi sono numeri naturali ed è possibile fornire regole di computazione per entrambi. In questo modo si riesce ad introdurre procedure induttive nello studio degli i-tipi e si dispone quindi di un mezzo potentissimo per la loro classificazione.

La nozione di rango permette di isolare varie classi di teorie dal punto di vista della realizzazione dei tipi, e in conseguenza di ciò della categoricità in potenza. Non è possibile qui entrare in dettagli ulteriori. Ci basti osservare che a partire dal 1965, anno in cui Morley presentò la sua teoria, si sono succeduti lavori molto importanti che sfruttano in molteplici direzioni il concetto di tipo e la possibilità di assegnare gradi. Tutti questi lavori sono culminati alla fine degli anni sessanta nei lavori di Saharon Shelah che ha gettato le basi per un'indagine sistematica di vastissimo raggio sul problema della classificazione e sulla *teoria della stabilità*. L'obietti-

vo di fondo è quello di fornire, per ogni possibile teoria elementare, teoremi di struttura e teoremi di non struttura, vale a dire trovare invarianti (cardinali, o altro) che caratterizzano i tipi di isomorfismo dei modelli, o provare che ciò è impossibile. Senza alcun dubbio il programma in questione è forse il tentativo più vasto e più ambizioso che sia stato formulato nella teoria dei modelli in questi vent'anni ed ha permesso di ottenere risultati di notevolissimo interesse.

Altra direzione di ricerca particolarmente viva negli anni sessanta e aperta da A. Robinson già negli anni cinquanta è la *model-theoretic algebra* il cui obiettivo è uno studio sistematico dell'algebra e della geometria algebrica utilizzando i metodi della teoria dei modelli. Le ricerche in questo campo hanno coinvolto in particolare concetti legati a quello di model-completezza, E.Q., completezza, ecc. mostrandone da una parte il significato algebrico (ad esempio la model-completezza della teoria dei campi algebricamente chiusi equivale al classico teorema di Hilbert noto come *Nullstellensatz*) dall'altra analizzandone i mutui rapporti. Così in particolare Shoenfield ha mostrato come la sub-completezza di cui sopra abbiamo parlato coincida con la eliminabilità dei quantificatori e ne fornisca quindi una versione semantica. Criteri particolarmente duttili per la E.Q. sono stati formulati da L. Blum, ancora Shoenfield ed altri, in connessione con proprietà algebriche fondamentali, quale l'amalgamazione, permettendo così di estendere i risultati per i campi algebricamente chiusi e reali chiusi ad altre classi di strutture, quali ad esempio i campi valutati, gli anelli ordinati, i campi differenziali, ecc.

I risultati ottenuti costituiscono la prova della fecondità del progetto originale di Robinson che era quello di vedere nella teoria dei modelli uno strumento per organizzare e sistemare concetti algebrici fondamentali, mostrando – e qui l'esempio della model-completezza è illuminante – come nozioni originariamente definite per strutture particolari, ammettessero una teoria generale che si applicava anche ad altre classi di strutture, permettendo così di rendere operanti analogie fra teorie diverse. L'interesse del concetto di model-completezza e di altre nozioni collegate che vedremo più avanti sta ad esempio nel fatto che apre la strada ad una teoria generale della proprietà «essere algebricamente chiuso», non più applicata ai soli campi, ma a strutture arbitrarie. Le inda-

gini sulla model-completezza sono così una chiave per capire quando, anche per queste altre strutture, è possibile una teoria delle strutture algebricamente chiuse.

Fatto importante, anche altre nozioni legate alla model-completezza hanno un significato algebrico; così ad esempio, come già intravisto da Tarski negli anni quaranta, il teorema di Chevalley, che stabilisce che la proiezione di una varietà algebrica su un campo algebricamente chiuso è una combinazione booleana di varietà algebriche di dimensione più bassa, è un corollario immediato della E.Q. per i campi algebricamente chiusi. Altro campo di indagine aperto da Robinson è quello che riguarda il diciassettesimo problema di Hilbert sulla rappresentabilità dei polinomi definiti positivi con coefficienti reali in termini di somme di frazioni razionali (quozienti di polinomi). Il problema era stato risolto da Artin nel 1927, ma Robinson negli anni cinquanta lo affrontò con i metodi della teoria dei modelli mostrandone il collegamento con la model-completezza della teoria dei campi reali chiusi. Fu questo risultato che dette inizio a interessanti ricerche sugli analoghi del problema per altri tipi di strutture.

Ancora più sorprendente è stato l'uso della E.Q. nello studio della geometria algebrica reale (ossia sul campo dei numeri reali). Qui l'approccio della *model-theoretic algebra* è stato insostituibile. In primo luogo, come mostrato da Jean Louis Krivine esso ha permesso la dimostrazione di analoghi del *Nullstellensatz* più generali di quelle fornite con metodi convenzionali da Dubois, in secondo luogo si è rivelato uno strumento centrale nel collegare la struttura *algebrica* delle varietà reali con quella *topologica* definita in base all'ordine, fornendo utilissimi criteri per determinare la natura topologica delle relazioni in esame. In tempi recenti la geometria algebrica reale su base modellistica ha raggiunto nuovi risultati (pensiamo in particolare alla definizione dello spettro di campi ordinati fornita da M. Coste) collegandosi con la logica categoriale ed è questo uno dei tanti esempi della fecondità di questo tipo di connubio. Altro campo di applicazione di questi metodi è l'algebra non commutativa, in particolare lo studio dei gruppi e dei corpi.

Un concetto introdotto ancora da Robinson negli anni sessanta e che in questo contesto si è rivelato molto fecondo è quello di *forcing* generalizzato cui abbiamo già accennato. Ci sono sostanzial-

mente due versioni del *forcing* nella teoria dei modelli: una versione finita, più raffinata, che è quella di cui si è servito A. MacIntyre negli studi sui gruppi algebricamente chiusi di cui sopra abbiamo parlato e quella infinita cui Robinson giunse nel 1970. L'idea di fondo è di concepire il *forcing* non più come relazione tra *insiemi* di formule e formule, ma tra *strutture* e formule. Supponiamo di avere una classe Σ di strutture *induttiva*, chiusa cioè rispetto alla somma su catene, e supponiamo che $\mathfrak{A} \in \Sigma$. La nozione $\mathfrak{A} \Vdash \mathscr{A}$ (la struttura $\mathfrak{A}$ forza la formula $\mathscr{A}$) si ottiene considerando le possibili estensioni di $\mathfrak{A}$ appartenenti a Σ. La definizione procede per induzione sulla complessità delle formule, e mentre nel caso di formule atomiche, di congiunzione, disgiunzione e quantificatore esistenziale coincide con la definizione di verità, se ne differenzia rispetto alla negazione: si definisce infatti che $\mathfrak{A} \Vdash \neg \mathscr{A}$ se non esiste estensione $\mathfrak{B}$ di $\mathfrak{A}$ in Σ tale che $\mathfrak{B} \Vdash \mathscr{A}$. Se definiamo $\forall x \mathscr{A}$ come $\neg \exists x \neg \mathscr{A}$ e $\mathscr{A} \supset \mathscr{B}$ come $\neg \mathscr{A} \vee \mathscr{B}$, avremo che anche l'implicazione e il quantificatore universale si comportano diversamente dal caso classico. Quella ora data è la definizione di *forcing* infinito; il caso finito si ottiene sostanzialmente prendendo come condizioni parti finite di diagrammi.

L'interesse della nozione di *forcing* sta nel fatto che consente di generalizzare il concetto di model-completezza. Chiamiamo infatti *generica* rispetto a una classe Σ una struttura $\mathfrak{A} \in \Sigma$ per cui verità e *forcing* coincidono: è facile dimostrare che se Σ è induttiva ogni struttura in Σ sarà sottostruttura di una struttura generica (questo corrisponde al teorema di Cohen sull'esistenza di condizioni complete di condizioni). Una struttura generica è una struttura in cui una formula $\mathscr{A}$ non può risultare falsa (non si può quindi avere $\mathfrak{A} \Vdash \neg \mathscr{A}$) se esiste almeno un'estensione di $\mathfrak{A}$ in cui $\mathscr{A}$ è vera, cioè in cui $\Vdash$ e verità coincidono. La cosa diviene interessante se $\mathscr{A}$ è una formula esistenziale contenente eventualmente nomi per elementi di $\mathfrak{A}$; dire che $\mathfrak{A}$ è generica significa in questo caso che se esistono elementi $x_1, \ldots, x_n$ in un'estensione di $\mathfrak{A}$ per cui $\mathscr{B}(a_1, \ldots, a_k, x_1, \ldots, x_n)$ è vera, allora elementi del genere esisteranno già in $\mathfrak{A}$ e ciò comporta che $\mathfrak{A}$ sarà *esistenzialmente chiusa* in Σ, cioè renderà vere tutte le formule esistenziali in $L_{\mathfrak{A}}$ vere in una qualche estensione di $\mathfrak{A}$ appartenente a Σ. In un certo senso quindi una struttura generica è una struttura con «sufficienti elementi».

Un esempio classico di strutture esistenzialmente chiuse è dato

dai campi algebricamente chiusi, come mostra il *Nullstellensatz* di Hilbert. Ogni formula $\exists y_1, ..., \exists y_n \mathscr{B}$ della teoria dei campi equivale ad una formula che afferma che esistono elementi $y_1, ..., y_n$ che soddisfano un opportuno sistema di uguaglianze e disuguaglianze fra polinomi, ed il *Nullstellensatz* afferma che se una formula del genere è soddisfatta in una estensione di un campo algebricamente chiuso lo è già nel campo stesso. C'è un altro fatto da notare: la chiusura esistenziale delle strutture generiche comporta che se $\mathfrak{A}$ e $\mathfrak{B}$ sono strutture generiche e $\mathfrak{A}$ è sottostruttura di $\mathfrak{B}$, $\mathfrak{A}$ sarà sottostruttura *elementare* di $\mathfrak{B}$. È quanto si verifica fra campi algebricamente chiusi di uguale caratteristica come dimostra la model-completezza della relativa teoria. Questo permette di capire in che senso il *forcing* consente di estendere il concetto di model-completezza. Supponiamo infatti che $\mathfrak{T}$ sia la teoria dei campi e $\mathscr{C}$ la classe dei suoi modelli; $\mathfrak{T}$ è chiaramente induttiva in quanto somme su catene di campi sono campi. Esiste quindi, per quanto detto sopra, una estensione generica per ogni campo $\mathfrak{A}$; d'altra parte, come abbiamo testé visto, se $\mathscr{C}$ è la classe dei campi generici e $\mathfrak{T}$ la sua teoria, $\mathfrak{T}$ sarà model-completa. Ne concludiamo allora che la teoria dei campi generici è la model-complezione della teoria dei campi. Ma Robinson aveva dimostrato che una teoria può avere al massimo una model-complezione e la model-complezione della teoria dei campi è la teoria dei campi algebricamente chiusi. Ne consegue che campi algebricamente chiusi e campi generici coincidono e che abbiamo trovato un nuovo modo per ottenere, quando esiste, la model-complezione di una teoria: basta considerare, nel caso la teoria sia induttiva, la classe delle strutture generiche e prenderne la teoria.

L'importanza di questo collegamento fra strutture generiche e model-complezione sta nel fatto che non sempre una teoria ha una model-complezione mentre è sempre possibile, purché sia induttiva, associarle il *forcing companion*: la teoria dei suoi modelli generici. Quel che conta è che quando la model-complezione esiste, essa coincide col *forcing companion*: si tratta quindi di una vera generalizzazione, la cui fecondità emerse chiaramente dopo che Paul Eklof e Gabriel Sabbagh dimostrarono nel 1971 che una teoria $\mathfrak{T}$ ha *model-companion* se e solo se le strutture esistenzialmente chiuse costituiscono una classe elementare. In combinazione con l'uso del *forcing*, questo permette di stabilire quando una teoria ha o

meno *model-companion*. Gli stessi Eklof e Sabbagh dimostrarono che la teoria dei gruppi non ha *model-companion* utilizzando questa tecnica, come fu fatto successivamente per i corpi, le algebre di Lie, gli anelli commutativi e i moduli su anelli non coerenti. Come ultimo tipo di collegamento che la tecnica del *forcing* ha realizzato, vogliamo citare gli sviluppi in termini categoriali che pongono in primo piano le strette analogie tra logica intuizionista, *forcing*, interpretazioni in prefasci e fasci. Un primo collegamento di questo tipo era stato posto in luce da MacIntyre nel 1973, ma fondamentali sono state le successive ricerche di Joyal del 1975.

Altro aspetto del *forcing* ampiamente studiato riguarda i collegamenti fra strutture generiche e altri tipi di strutture «con sufficienti elementi». I risultati più significativi riguardano la possibilità di caratterizzare la classe delle strutture generiche rispetto a una data Σ in termini che non coinvolgono il *forcing* e l'elaborazione di metodi per la costruzione di strutture generiche. È questo un campo particolarmente sviluppato negli anni sessanta e i risultati vanno da quelli di Harry Simmons che hanno permesso di approssimare «dall'alto» le classi delle generiche a quelli di Mitsuru Yasuhara che hanno collegato strutture sature e strutture generiche. Un problema interessante è quello di vedere entro che limiti il *forcing* di Cohen si possa presentare come caso particolare di quello di Robinson, e Keisler nel 1973 ha fatto alcuni passi decisivi in questa direzione, mostrando il collegamento fra i due tipi di *forcing* e i teoremi di omissione di tipo. Come molte delle tecniche della teoria dei modelli classica, il *forcing* si può estendere a linguaggi più forti di quelli elementari, e il collegamento è interessante in quanto esistono risultati (ad esempio di Carol Wood) che mostrano come sia possibile assiomatizzare classi di strutture generiche utilizzando linguaggi infinitari.

Il discorso su questi linguaggi è d'altro canto uno dei temi centrali della ricerca degli anni sessanta e ha registrato notevolissimi sviluppi in più direzioni. Il problema generale che si pone nei confronti dei linguaggi infinitari (e delle altre generalizzazioni dei linguaggi elementari) è quello di vedere entro che limiti si estendono ad essi risultati validi per i linguaggi elementari. Anche se esistono tecniche specifiche per i linguaggi infinitari è chiaro infatti che per avere una teoria dei modelli significativa occorre che almeno una parte dei risultati sui linguaggi elementari si possa

estendere. Di fatto, si verifica che solo gli estremi dello spettro di questi linguaggi ammettono una teoria dei modelli siffatta. Abbiamo così una teoria estremamente ricca per i linguaggi infinitari più deboli, del tipo $L_{\omega_1,\omega}$ e all'estremo opposto per i linguaggi $L_{\alpha,\beta}$ dove α e β sono particolari «grandi» cardinali. La situazione cambia considerando opportuni frammenti e superando considerazioni semplicemente cardinali. Questo dipende dal fatto, posto in luce con evidenza dal caso degli insiemi ammissibili, che per sviluppare una teoria dei modelli significativa sono necessari in qualche modo «principi di riflessione» sui cardinali che si ammettono. Il discorso sui linguaggi infinitari si può applicare anche ad altri tipi di estensione dei linguaggi elementari, quali le varie formulazioni di ordine superiore, i linguaggi con quantificatori generalizzati, ecc.

Il teorema centrale della teoria dei modelli per linguaggi elementari, che nella maggior parte dei casi viene meno per linguaggi infinitari (salvo opportune restrizioni come quella di Barwise di cui abbiamo parlato a suo tempo per i frammenti ammissibili) è il teorema di compattezza, e questo ha fatto sì che ci sia stato un intergioco molto stretto fra studio dei grandi cardinali e teoria di questi linguaggi, cercando per quali cardinali α e β il teorema di compattezza risulta valido. Così è stato introdotto il concetto di cardinale *compatto* (fortemente o debolmente) che *grosso modo* è appunto un cardinale per cui, considerando un linguaggio $L_{\alpha,\beta}$ con $\alpha \geqslant \beta$, si ha che vale un teorema di compattezza che afferma che ogni insieme X di formule ha un modello se e solo se lo ha ogni sua parte di cardinalità $< \alpha$. Il concetto, introdotto da Tarski, è stato ampiamente studiato da Tarski stesso in collaborazione con Keisler nel citato lavoro del 1964, e quindi da Hanf e Scott tra gli altri. La conclusione è che, se esistono, i cardinali compatti sono straordinariamente grandi. Come infatti ha dimostrato Hanf, tutti i cardinali inaccessibili $> \omega$ sono fortemente non compatti, con la possibile eccezione di alcuni iperinaccessibili.

Non è possibile qui dilungarci sugli studi condotti in questo campo. Ci basti dire che mentre il teorema inferiore di Löwenheim-Skolem – fatte opportune precisazioni – vale in generale, la situazione è più complessa nel caso della versione superiore. In questo contesto Hanf ha dimostrato l'esistenza dei numeri di

Hanf, di cui abbiamo già parlato, per classi di linguaggi molto generali e l'attenzione dei ricercatori si è concentrata sul computo di tali cardinali nei casi specifici. Se così in generale uno dei metodi fondamentali per costruire modelli – quello di Henkin – viene a cadere, altri metodi, quale quello degli ultraprodotti, si possono in generale estendere con modificazioni che riguardano la natura degli ultrafiltri considerati.

Ma è soprattutto la tecnica degli *isomorfismi locali* introdotta da Roland Fraïssé negli anni cinquanta che si è rivelata uno degli strumenti più importanti nello studio dei rapporti fra strutture dal punto di vista dei linguaggi infinitari. Questo è emerso in particolare in seguito ai lavori di C. Karp, di David Kueker ed altri che hanno posto in luce lo stretto rapporto fra i tipi di prolungamenti di isomorfismi locali che si considerano e la lunghezza di successioni di variabili che si possono quantificare. L'idea che sta alla base della tecnica di Fraïssé risale alla dimostrazione cantoriana dell'isomorfismo degli insiemi densi numerabili senza estremi. Cantor costruiva l'isomorfismo «incollando» isomorfismi parziali di sottostrutture delle strutture date. La nozione di isomorfismo locale si ottiene concentrandosi appunto su questi isomorfismi parziali.

Consideriamo due strutture $\mathfrak{A}$ e $\mathfrak{B}$ e sia H la famiglia di tutti gli isomorfismi fra sottostrutture finitamente generate di $\mathfrak{A}$ e sottostrutture finitamente generate di $\mathfrak{B}$. L'idea di Fraïssé è di analizzare l'equivalenza fra due strutture rispetto alle proprietà esprimibili in dati linguaggi utilizzando isomorfismi locali e loro prolungamenti. Limitiamoci al caso dei linguaggi elementari e diciamo che due strutture $\mathfrak{A}$ e $\mathfrak{B}$ sono parzialmente isomorfe quando esiste un sottoinsieme non vuoto F di H che gode delle due proprietà di prolungamento seguenti (*back and forth*):

B) per ogni $f \in F$ ed ogni $a \in \mathfrak{A}$ esiste in F un isomorfismo f' che estende f ed ha a nel suo dominio;

F) per ogni $f \in F$ ed ogni $b \in \mathfrak{B}$ esiste in F un isomorfismo f' che estende f e che ha b nel suo codominio.

È immediato verificare a questo punto che se $\mathfrak{A}$ e $\mathfrak{B}$ sono parzialmente isomorfe, saranno elementarmente equivalenti. Questo

si verifica per induzione sulla complessità delle formule e dipende dal fatto che se due strutture sono isomorfe rendono vere le stesse formule, cosicché nel caso di formule senza quantificatori è chiaro che se $a_1, ..., a_n$ sono elementi di $\mathfrak{A}$ essi soddisferanno una formula $\mathscr{D}$ priva di quantificatori se e solo se esiste un isomorfismo $f \in F$ che ci dà degli elementi $b_1,...,b_n$ appartenenti a $\mathfrak{B}$ che soddisfano $\mathscr{D}$ in $\mathfrak{B}$. Che questo isomorfismo esista è ovvio considerando che 1) $a_1, ..., a_n$ individuano una sottostruttura $\mathfrak{A}_0$ finitamente generata di $\mathfrak{A}$ e 2) che applicando n volte la clausola B) avremo l'isomorfismo parziale richiesto.

Il ricorso ai prolungamenti interviene per valutare i quantificatori. Consideriamo ad esempio la formula $\exists x \mathscr{D}$ e supponiamo che $\mathfrak{A} \Vdash \exists x \mathscr{D}(a_1, ..., a_n)$. Esisterà allora un $a \in \mathfrak{A}$ per cui $\mathfrak{A} \Vdash \mathscr{D}(a, a_1, ..., a_n)$. Se assumiamo per ipotesi che $f \in F$ sia un isomorfismo della sottostruttura generata da $a_1, ..., a_n$ in una sottostruttura di $\mathfrak{B}$, la clausola B) ci garantisce che esso si potrà prolungare in un isomorfismo f' che ha anche a nel suo dominio. Questo ci permette di concludere che $\exists x \mathscr{D}$ sarà soddisfatta anche in $\mathfrak{B}$ dalle immagini degli a_i. L'isomorfismo locale ci dà quindi una condizione sufficiente per concludere all'equivalenza elementare. Ma non vale in generale il viceversa e per avere una caratterizzazione dell'equivalenza elementare in termini di isomorfismi locali occorre – come fece Fraïssé – raffinare l'analisi e considerare famiglie incapsulate l'una nell'altra di isomorfismi parziali, tali che ogni isomorfismo di un insieme della famiglia si prolunga in un isomorfismo appartenente all'insieme successivo. In questo modo se ci limitiamo a linguaggi senza costanti funzionali e finiti, è possibile determinare a meno di isomorfismi ogni sottostruttura finitamente generata (quindi, non essendoci costanti funzionali, finita) mediante una formula, e possiamo così concludere, dalla equivalenza elementare, l'esistenza della famiglia di isomorfismi. Ad un'analoga caratterizzazione dell'equivalenza elementare utilizzando isomorfismi parziali giungeva anche Ehrenfeucht nel 1961 ma quel che conta è che più che per i linguaggi elementari, il metodo degli isomorfismi locali si mostra particolarmente illuminante nel caso dei linguaggi infinitari una volta che: 1) si sostituisca la condizione sulla generabilità *finita* delle sottostrutture in esame con la vincolante, ad esempio, di avere cardinalità κ (con κ cardinale infinito) e si assuma eventualmente che è possibile «incollare» catene di iso-

morfismi locali di cardinalità minore di κ, 2) si sostituisca nelle clausole B) e F) la prolungabilità un elemento per volta con la λ-prolungabilità, per cui ad esempio la clausola B) diverrà

Bλ) per ogni $f \in F$ ed ogni $X \subseteq \mathfrak{B}$ tale che $\overline{\overline{X}} < \kappa$ esiste un $g \in F$ che estende f e il cui codominio include X.

Mentre la clausola sulla cardinalità delle sottostrutture è legata alla possibilità di operare nel linguaggio meno di κ congiunzioni e disgiunzioni, le condizioni sulla prolungabilità riflettono la possibilità di quantificare simultaneamente meno di λ variabili cosicché emerge uno strettissimo legame fra caratterizzabilità utilizzando linguaggi infinitari e metodo di Fraïssé modificato come sopra. Negli anni sessanta la Karp riuscì così a dimostrare che, se chiamiamo $L_{\infty,\omega}$ ogni linguaggio in cui si possono fare congiunzioni e disgiunzioni arbitrarie ma si possono quantificare solo un numero *finito* di variabili per volta, si ha che la equivalenza $\equiv_{\infty,\omega}$ rispetto a formule di questi linguaggi è *completamente* caratterizzabile in termini di isomorfismi locali.

Questi risultati si possono raffinare considerando i linguaggi $L_{\kappa,\lambda}$ come è stato fatto da diversi ricercatori, fra cui in particolare Eklof e Barwise nei primi anni settanta, Kueker alla fine degli anni sessanta, ecc. In questo modo non solo si ottengono criteri sufficienti per la equivalenza rispetto ai linguaggi in oggetto ma anche condizioni di immergibilità di una struttura in un'altra mediante monomorfismi, collegamenti fra le proprietà dei fattori e dei prodotti di strutture, condizioni di isomorfismo, ecc. Tutti questi risultati dipendono dalla natura particolare dei cardinali κ e λ cosicché ad esempio si ottiene il teorema di Cantor sull'isomorfismo fra ordini densi senza estremi numerabili, sfruttando il fatto che catene di isomorfismi di dominio finito indiciate dai naturali hanno somme che stabiliscono un isomorfismo tra le riunioni dei domini e dei codomini. Così ancora si può provare la generalizzazione del teorema di Cantor data da Hausdorff sugli η_λ-insiemi – cioè quegli insiemi linearmente ordinati – per cui data una sezione (A, B) dove tanto A che B hanno cardinalità minore di $\aleph_\lambda$ esistono elementi x, y, z, tali che per ogni $a \in A$ e ogni $b \in B$ si abbia

$$x < a < y < b < z.$$

Se $\lambda = 0$ gli η_λ-insiemi coincidono con gli ordini densi senza estremi e il teorema di Hausdorff generalizza quello di Cantor nel senso che afferma che se λ è regolare e infinito due qualunque di questi insiemi sono isomorfi. Con la stessa tecnica si dimostrano i vari risultati sulla unicità a meno di isomorfismi tra modelli atomici elementarmente equivalenti e sulle strutture, sempre elementarmente equivalenti, che godono di un «opportuno» grado di saturazione: prima si prova che data la natura particolare delle strutture l'equivalenza elementare comporta un opportuno tipo di isomorfia locale, quindi si «incollano» gli isomorfismi locali di cui si è così provata l'esistenza per ottenere un isomorfismo fra le strutture.

Col tempo le tecniche di Fraïssé e Ehrenfeucht si sono andate sempre più rivelando uno strumento centrale non solo nello studio dei linguaggi infinitari ma anche di altre estensioni dei linguaggi elementari quali quelli con quantificatori generalizzati, forme deboli della logica del secondo ordine, ecc. È proprio questo studio che ha permesso significative applicazioni a problemi di carattere algebrico riguardanti ad esempio i gruppi abeliani e la generalizzazione del teorema di Ulm (fondamentali in questo senso alcuni lavori di Barwise e Eklof dei primi anni settanta) la costruzione di una teoria dei modelli per strutture topologiche (dovuta a T.A. McKee, Martin Ziegler, J. Flum, ed altri) ecc.

Queste estensioni portano naturalmente ad un problema più generale: quello di determinare delle condizioni sufficienti e necessarie affinché un linguaggio abbia una «buona» teoria dei modelli, il che significa sostanzialmente chiedersi quante delle proprietà dei linguaggi elementari valgono per essi e sotto quali condizioni. La questione acquista un significato più generale se si pensa che i linguaggi elementari, come ha mostrato la pratica logica e non solo logica di almeno mezzo secolo, costituiscono un vero e proprio paradigma nella determinazione del concetto stesso di linguaggio logico. Quello che ci si domanda allora è come estrarre le caratteristiche generali del concetto di logica analizzando i linguaggi elementari. I primi risultati in questo senso furono ottenuti da Per Lindström nel 1969, ma per un certo tempo non ricevettero l'attenzione che meritavano. È stato solo verso il 1971 che da più parti si è riconosciuta la portata di questi lavori ed è iniziata una ricerca di vasto raggio in quella che oggi si chiama «logica astratta» o *soft model theory*.

Il programma di Lindström è sostanzialmente quello di caratterizzare la logica dei linguaggi elementari all'interno delle sue possibili estensioni. Per far questo è necessario disporre di una nozione generale di logica, la logica astratta appunto, entro la quale isolare quella elementare mediante un'opportuna famiglia di proprietà. Una logica astratta (una **LA**) **L** è determinata dall'insieme S dei suoi enunciati e dalla sua relazione di verità $\models$; quest'ultima relazione è definita fra strutture ed enunciati di **L**. Possiamo quindi identificare una logica astratta **L** con la coppia $< S, \models >$. Perché abbia senso parlare di logica occorre porre delle condizioni sulla relazione di verità e sull'insieme S. Sostanzialmente queste condizioni sono che la verità è preservata dall'isomorfismo, che ogni enunciato ha un vocabolario extralogico finito, che è possibile estendere opportunamente una struttura ad un'altra con linguaggio più ricco, ed infine che gli enunciati di S sono chiusi rispetto a congiunzione, negazione e a un'opportuna nozione di sostituzione. Chiaramente, la logica elementare **LE** è una **LA** e tra **LA** è possibile definire diverse relazioni d'ordine basandosi sulla capacità espressiva dei relativi linguaggi. Possiamo così dire che $\mathbf{L} \subseteq \mathbf{L}'$ se e solo se per ogni enunciato $\mathscr{A}$ di **L** c'è un enunciato $\mathscr{B}$ di $\mathbf{L}'$ tale che $\mathscr{A}$ e $\mathscr{B}$ hanno gli stessi modelli. Questo significa che come caratteristica dei linguaggi si pone la loro capacità di individuare classi di strutture e la maggior potenza dei linguaggi è data dalla maggiore capacità espressiva. La relazione $\subseteq$ si può raffinare considerando classi opportune di strutture. Possiamo così definire che $\mathbf{L} \subseteq_{inf} \mathbf{L}'$ se e solo se per ogni enunciato $\mathscr{A}$ di **L** c'è un enunciato $\mathscr{B}$ di $\mathbf{L}'$ tale che $\mathscr{A}$ e $\mathscr{B}$ hanno gli stessi modelli infiniti. Tanto a partire da $\subseteq$ che da $\subseteq_{inf}$ possiamo chiaramente definire due relazioni di equivalenza $\equiv$ e $\equiv_{inf}$ ponendo che $\mathbf{L} \equiv \mathbf{L}'$ se e solo se $\mathbf{L} \subseteq \mathbf{L}'$ e $\mathbf{L}' \subseteq \mathbf{L}$ e analogamente per $\equiv_{inf}$.

A questo punto quello che ci si domanda è come caratterizzare **LE** all'interno della famiglia delle **LA**. L'idea è di utilizzare i teoremi fondamentali della teoria dei modelli di **LE** come altrettante proprietà caratteristiche e di indagare quali sono le estensioni di **LE** che godono di tutte o di parte di queste proprietà. Alcune delle proprietà centrali isolate da Lindström sono la proprietà di compattezza, PC (ogni insieme di enunciati T ha un modello se e solo se ogni suo sottoinsieme finito possiede un modello), la proprietà di Löwenheim, PL (ogni enunciato di **L** se ha un modello infinito

ne ha uno numerabile), la proprietà di Tarski, PT (se un enunciato $\mathscr{A}$ ha un modello infinito ne ha uno più che numerabile). Tutte queste tre proprietà sono chiaramente godute da **LE** e corrispondono a teoremi classici. Un primo risultato ottenuto da Lindström stabilisce che non esiste $\mathbf{L} \supseteq \mathbf{LE}$ e non equivalente a **LE** che goda di PL e PC rispetto agli insiemi T numerabili: queste due proprietà sono quindi sufficienti a caratterizzare **LE** nella famiglia delle sue estensioni. Un'altra caratterizzazione fa ricorso alla proprietà dell'unione di catene, PUC, che stabilisce che se $\mathscr{K}$ è una catena di strutture rispetto all'estensione elementare relativa alla logica in questione, allora la somma su questa catena è estensione elementare di ogni elemento di $\mathscr{K}$; Lindström dimostra che se $\mathbf{L} \supseteq \mathbf{LE}$ e **L** gode di PC e di PUC, allora $\mathbf{L} \equiv \mathbf{LE}$.

Questi sono solo alcuni esempi delle diverse caratterizzazioni che Lindström ha ottenuto; altre se ne possono avere considerando combinazioni di ulteriori proprietà, quali ad esempio la proprietà di completezza, delle tavole semantiche, dell'interpolazione ecc. e la ricerca in questo campo continua tuttoggi ad opera, ad esempio, di S. Shelah, H. Friedman, Johann Makowski, Daniele Mundici, ecc. Il senso di questi risultati è duplice: da una parte ci mostrano come, data una definizione sufficientemente generale di logica, l'essenza della logica elementare stia in una certa combinazione di proprietà che riguardano le classi di modelli che in essa è possibile determinare e quindi costituiscono una precisazione dell'idea dei linguaggi come «crivelli» per la classificazione di strutture; dall'altra essi costituiscono un effettivo strumento nella ricerca di estensioni di **LE** che abbiano una «buona» teoria dei modelli, che siano plausibili come logiche. Essi ci dicono infatti quando è possibile individuare delle estensioni proprie di **LE**, essenzialmente più potenti, che godano di certe combinazioni di proprietà, dandoci così un metodo generale per stabilire quando la ricerca di date logiche è possibile o meno.

La nozione di logica astratta data da Lindström non è però la più generale possibile; la restrizione riguardante la *finitezza* delle costanti extralogiche in ogni formula esclude ad esempio i linguaggi infinitari, prendendone in considerazione solo frammenti. Estendendo opportunamente questa definizione è possibile ottenere ulteriori risultati e tentare anche la caratterizzazione di logiche diverse da quella elementare. La candidata più immediata per

questa indagine è la logica $\mathbf{L}_{\omega_1,\omega}$ che ammette congiunzioni e disgiunzioni infinite numerabili ma quantificazioni solo finite. Nel 1971 Barwise è riuscito ad ottenere risultati riguardanti la sua caratterizzazione all'interno di una famiglia di logiche molto generali. Va sottolineato che considerando linguaggi come $L_{\omega_1,\omega}$ si pongono dei problemi nella scelta delle estensioni che non si presentano nel caso finitario. Sappiamo ad esempio che in **LE** tutti i connettivi sono definibili a partire da $\neg$ e da $\wedge$, sicché la scelta di questi tra gli altri connettivi non comporta problemi. Se si ammettono però congiunzioni infinitarie e quantificazioni finite non si dispone più di teoremi tipo quelli della forma normale prenessa; visto il ruolo che queste forme hanno nello studio dei modelli si pone naturale il problema di che altri connettivi vadano introdotti. Risultati di caratterizzazione quali quelli ottenuti da Barwise danno un filo conduttore nella scelta delle varie possibilità, permettendo così di individuare criteri di «naturalezza» nella scelta dei concetti logici fondamentali.

Un altro aspetto del lavoro di Barwise riguarda il problema dell'assolutezza (nel senso di Gödel) della logica adottata, l'indipendenza cioè di sue proprietà specifiche dall'universo degli insiemi in cui la studiamo. In questo contesto astratto in cui le formule possono essere oggetti *infiniti*, l'insieme degli enunciati di un linguaggio o la relazione di verità dipendono dall'universo in cui ci poniamo, estendendolo o restringendolo possono cambiare gli aspetti caratteristici della logica. Questo non avviene se, adottata come metateoria una data teoria $\mathfrak{T}$ degli insiemi, poniamo delle condizioni sulla forma dei predicati in $\mathfrak{T}$ che definiscono l'insieme delle formule e la relazione di verità per la logica in esame. Sul peso determinante del tipo di definibilità insiemistica della sintassi di linguaggi infinitari in esame ci siamo già soffermati parlando degli insiemi ammissibili, dei frammenti che essi determinano e della possibilità di stabilire per essi risultati tipo il teorema di compattezza di Barwise.

Qui il problema è più generale e riguarda l'indipendenza della logica che studiamo dall'universo insiemistico in cui ci si pone. Una logica astratta come proposto da Barwise è così *assoluta* relativamente a una teoria $\mathfrak{T}$ se la relazione di verità è espressa da un predicato di tipo $\Delta_0^{\mathfrak{T}}$ e l'insieme delle formule da un predicato $\Sigma_1^{\mathfrak{T}}$. Ciò significa, come sappiamo da quanto detto a suo tempo sugli

insiemi, che mentre la proprietà di essere formula sarà persistente, quella di essere vero sarà assoluta. È importante notare che sarebbe troppo restrittivo assumere nella definizione di logica assoluta che anche la proprietà di essere formula sia $\Delta_0^{\mathfrak{T}}$; sappiamo infatti che la proprietà di essere numerabile, e quindi quella di essere congiunzione numerabile di formule, non è assoluta. Queste logiche generali sono state oggetto di ricerca ad opera di vari studiosi tra i quali Friedman, Makowski, Feferman, ecc. In questa linea lo stesso Barwise ha tentato nel 1974 di giungere a una formulazione di una *teoria assiomatica dei modelli* basandosi su una definizione astratta di linguaggio e di verità. I risultati ottenuti testimoniano un interesse che sempre più si sta imponendo nella ricerca: l'indagine del concetto generale di logica e lo studio del ruolo che l'individuazione dell'articolazione logica ha sull'analisi degli oggetti matematici (e no).

3. APPLICAZIONI DELLA TEORIA DEI MODELLI

Abbiamo lasciato per ultimo un tipo di indagine che – sorto all'interno della teoria dei modelli – mostra con particolare evidenza come la ricerca logica possa, in alcuni casi, volgersi dall'analisi di quello che il matematico fa *concretamente*, creando nuovi tipi di oggetti che permettono un rapporto diverso tra la matematica e i suoi campi d'applicazione, un nuovo modello di *matematizzazione della realtà*.

Ci riferiamo alla matematica non standard ed in particolare alla aritmetica e all'Analisi non standard. Se già dai primi anni sessanta – dopo i primi studi di Robinson sull'argomento – era risultato chiaro come i metodi introdotti fossero in grado, se non altro, di fornire uno sviluppo alternativo a capitoli centrali dell'Analisi intuitivamente e pedagogicamente significativo oltre che fecondo sul piano della ricerca, la storia dell'aritmetica non standard è stata ben diversa e si è presentata – per usare le parole di Robinson – come un «deserto desolato» rispetto alla ricchezza di applicazioni della teoria dei modelli ad algebra ed Analisi.

Sostanzialmente – come abbiamo ricordato a suo tempo – i primi modelli non standard dell'aritmetica (vale a dire le prime strutture modello di $\mathfrak{P}$ non isomorfe ad N) erano stati presentati

da Skolem nel 1933, ma solo trent'anni dopo le ricerche di Skolem erano divenute accessibili ad un pubblico matematico più ampio. Curiosamente ciò avvenne in un articolo di una raccolta, *Mathematical Interpretation of Formal Systems* (*Interpretazione matematica di sistemi formali*, 1955) in cui compariva anche l'articolo di Łos nel quale veniva presentata l'operazione di ultraprodotto. Il fatto è curioso in quanto, come non si tardò a realizzare, la costruzione di Skolem è un tipo particolare di ultraprodotto.

Per Skolem i modelli non standard erano essenzialmente solo una prova di come non avessero determinazione univoca i concetti fondati su base insiemistica e in quanto tali per lui non avevano un interesse matematico specifico. Lo stesso discorso si può fare per tutta un'altra famiglia di modelli non standard dell'aritmetica di Peano, la cui esistenza scende dal teorema di incompletezza di Gödel. Si è soliti distinguere tra questi ultimi modelli non standard che non sono isomorfi ad N, ma non sono neppure ad essa elementarmente equivalenti, e quelli che sono estensioni elementari di N ma non sono ad essa isomorfi: i secondi vengono chiamati modelli non standard *forti*, mentre i primi sono semplicemente i modelli non standard o non standard *deboli*.

Già nei primi anni cinquanta Henkin aveva mostrato come fosse facile costruire modelli non standard forti sfruttando la compattezza. Se infatti T è l'insieme di tutti gli enunciati veri in N, e c una costante che non occorre in T, consideriamo l'insieme

$$T \cup \{c \neq \bar{n} \mid n \in \mathrm{N}\}$$

dove $\bar{n}$ è la cifra per il numero n.

È facile verificare che ogni parte finita X di questo insieme avrà un modello (basta infatti considerare la stessa struttura N ed interpretare la costante c su un qualunque naturale maggiore di quelli la cui cifra occorra nel frammento X dell'insieme). Per il teorema di compattezza l'intero insieme avrà un modello N^* che conterrà un elemento almeno – il denotato di c – che non coinciderà con nessun naturale. N^* sarà così un'estensione *propria* di N, linearmente ordinata ma *non* bene ordinata. Come Henkin dimostrava, il tipo d'ordine di N^* è

$$\omega + (\omega^* + \omega)\eta$$

dove η è il tipo d'ordine dei razionali. In altre parole l'ordine di N* si ottiene prendendo N come segmento iniziale ed aggiungendo tante copie degli interi (il cui tipo d'ordine è $\omega^* + \omega$) quanti sono i razionali. Henkin dimostrava più in generale che *ogni* modello non standard dell'aritmetica doveva avere tipo d'ordine come quello indicato salvo che al posto di η poteva comparire qualsiasi tipo d'ordine denso.

Questo fatto dava già una prima misura della struttura «bizzarra» dei modelli non standard, e testimoniava le difficoltà di un loro studio sistematico. La cosa importante – il cui pieno significato sarebbe emerso solo più tardi, dopo la creazione dell'Analisi non standard – riguardava la *natura* degli elementi dei modelli che non erano numeri naturali. Come indicava il risultato di Henkin, nell'ordine di ogni modello essi si situano *dopo* i naturali, così che si possono concepire come «naturali infiniti», maggiori di ogni naturale finito, e fu proprio in questo loro ruolo che vennero utilizzati negli anni cinquanta nello studio dell'aritmetica: la prima volta da Ryll-Nardzewski per dimostrare la non assiomatizzabilità finita di $\mathfrak{P}$, la seconda da M. Rabin nel 1962, che estese il risultato precedente dimostrando che $\mathfrak{P}$ è essenzialmente illimitata nel senso che nessuna sua estensione coerente può essere data da assiomi i cui prefissi hanno complessità limitata fissata a priori. Questo risultato si collegava a ricerche più o meno coeve di Kreisel, Feferman, Scott e S. Tennenbaum sull'esistenza di teorie finitamente assiomatizzabili senza modelli ricorsivi, permettendo tra l'altro di fornire una nuova dimostrazione del teorema di Mostowski secondo il quale l'insieme degli enunciati validi in tutti i modelli ricorsivi non è un insieme aritmetico (nel senso che non appartiene alla gerarchia aritmetica).

Tutte queste applicazioni riguardavano però non i modelli non standard in sé, ma il loro uso nello studio delle teorie. Questo, si badi, anche se Rabin, negli stessi anni, aveva dimostrato fatti che avevano i modelli in sé come centro dell'attenzione, quali ad esempio l'esistenza per ogni modello non standard di una sua estensione propria in cui risultavano risolubili equazioni diofantee non risolubili nella struttura di partenza. L'idea di Rabin era la stessa che Robinson aveva ampiamente illustrato nel 1961 nel suo contributo al volume *Infinitistic Methods* (*Metodi infinitari*) in cui poneva come programma l'indagine sistematica dei modelli non

standard dell'aritmetica come metodo per affrontare specifici problemi della teoria dei numeri. Di fatto sarà solo dopo lo sviluppo dell'Analisi non standard che questo programma comincerà a prendere forma anche se già nei primi anni sessanta si ebbero interessanti contributi di Scott e Mendelson, sulla struttura algebrica dei modelli non standard, e di R. MacDowell e Specker sulla struttura d'ordine di questi modelli, in particolar modo sulle estensioni finali (le estensioni $\mathfrak{M}'$ di un modello $\mathfrak{M}$ in cui gli elementi nuovi vengono tutti dopo gli elementi vecchi già appartenenti ad $\mathfrak{M}$).

Il parallelo con l'Analisi non standard mostra il ruolo centrale dei naturali infiniti. Il fatto fondamentale, provato da Robinson, è il cosiddetto «principio del traboccamento» (*Overspill*) secondo il quale se $\mathfrak{M}$ è un modello non standard e d un suo elemento, data una formula $\mathscr{A}$ con due sole variabili libere

$$\forall x \in \omega \, (\mathfrak{M} \models \mathscr{A}(x, \bar{d})) \text{ se e solo se}$$
$$\exists a \text{ infinito } (\mathfrak{M} \models \forall x < \bar{a} \, \mathscr{A}(x, \bar{d})).$$

Il principio si può rafforzare, ma mostra bene come la validità di un fatto riguardante *tutti* i naturali si possa ridurre alla esistenza di naturali infiniti con particolari proprietà. Come sottolineato da C. Smorynski, i ruoli che i naturali infiniti possono svolgere sono almeno due: uno – posto in luce dal principio sopra citato – è quello di mostrare come si comportano i naturali che un numero infinito incontra sul suo cammino; il secondo – che emerge una volta che si facciano interagire numeri infiniti e tecniche di aritmetizzazione – è quello di permettere la codifica di insiemi *infiniti* di numeri naturali, consentendo di associare ad ogni modello un sistema di insiemi di numeri naturali – il suo *sistema standard* analizzato da Scott già nel 1962 – che riflette proprietà centrali del modello in esame.

I risultati più interessanti riguardanti i modelli non standard nascono dall'intergioco fra questi ruoli, che ha mostrato come sia possibile affrontare questioni classiche di incompletezza e indecidibilità rimpiazzando costruzioni per diagonalizzazione con la costruzione di modelli non standard. Così ad esempio Robinson ha potuto provare il teorema di Tarski sulla non definibilità della verità aritmetica ricorrendo ai modelli non standard e in

tempi più vicini a noi H. Friedman ha potuto analizzare problemi di immergibilità fra modelli non standard in termini dei sistemi standard associati. Lo sviluppo più ricco di queste indagini fuoriesce dai limiti temporali che ci siamo assegnati, ed è in larga misura collegato ai risultati di indecidibilità tipo quelli di Paris e Harrington e alle ricerche iniziate da Friedman nel 1973 sui modelli non standard non solo dell'aritmetica ma della stessa teoria degli insiemi.

L'altro fatto che dà un sapore particolare allo studio dei modelli non standard dell'aritmetica è il rapporto strettissimo tra considerazioni semantiche e temi propri della teoria della ricorsività su cui già sopra abbiamo richiamato l'attenzione: uno degli aspetti di questa interazione è il rilievo dato in questo contesto ai modelli *ricorsivamente saturi*, il cui studio è stato iniziato da Donald Jensen e Ehrenfeucht. Il concetto di modello ricorsivamente saturo per teorie generali era stato isolato da J.P. Ressayre e J. Schlipf e i primi lavori significativi sull'argomento risalgono alla metà degli anni settanta. La fecondità di questa nozione nasce dal fatto che, per definizione, perché un modello sia ricorsivamente saturo basta che siano realizzati quei tipi che costituiscono insiemi ricorsivi e che quindi, se non finiti, siano definibili mediante *schemi* finiti.

Come sappiamo è spesso complicato dimostrare l'esistenza di modelli saturi per una teoria e questo spesso per ragioni puramente insiemistiche e quindi – in certa misura – che non riguardano le proprietà logiche delle teorie in esame. Il problema è molto più semplice se ci limitiamo a modelli ricorsivamente saturi e si può dimostrare che ogni teoria coerente in un linguaggio numerabile ha modelli finiti o ricorsivamente saturi di cardinalità numerabile. Questo fa sì che i modelli in questione risultino estremamente utili per fornire dimostrazioni semantiche di teoremi sintattici (un esempio per tutti: il teorema di interpolazione) senza dover ricorrere ai modelli saturi. L'interesse dei modelli ricorsivamente saturi nello studio dell'aritmetica è più specificamente legato al ruolo particolare che hanno considerazioni di effettività nel caso dei modelli della teoria $\mathfrak{P}$. Altri esempi di questo intergioco riguardano il ruolo del teorema di Matijasievič dal punto di vista semantico nel permettere una caratterizzazione delle relazioni ricorsive che coinvolge non solo i modelli stan-

dard, ma anche quelli non standard; come già abbiamo ricordato, a questi argomenti sono dedicati alcuni importanti lavori di H. Gaifman, il quale, tra l'altro, ha dimostrato una sorta di teorema di forma normale per i modelli non standard analogo a quello di Steinitz per le estensioni dei campi, in base al quale ogni estensione di un modello $\mathfrak{M}$ di $\mathfrak{P}$ si può ottenere in due passi: prima mediante un'estensione elementare *cofinale* $\mathfrak{M}'$ (tale cioè che ogni elemento di $\mathfrak{M}$ è maggiorato da un elemento di $\mathfrak{M}'$) poi prendendo un'estensione finale $\mathfrak{M}''$ di $\mathfrak{M}'$.

Se lo studio dei modelli non standard dell'aritmetica ha avuto il suo massimo sviluppo dopo la metà degli anni settanta, l'Analisi non standard aveva rivelato le sue potenzialità sin dal suo apparire e numerosi sono stati non solo i logici, ma i matematici tradizionali che hanno affiancato Robinson nel sistematico impiego di metodi non standard nella topologia generale, nell'Analisi funzionale, nella teoria della probabilità, nella fisica, nell'economia matematica, ecc. Nel 1966 usciva un lavoro congiunto di Robinson e di A. Bernstein in cui si offriva una soluzione ad un problema posto da K.T. Smith e P.R. Halmos nel contesto dell'Analisi funzionale e che fornisce forse il primo esempio di risultato matematico ottenuto in prima istanza con metodi non standard. Nello stesso numero della rivista, all'articolo di Robinson e Bernstein seguiva un articolo dello stesso Halmos in cui il risultato veniva ridimostrato con metodi standard. Di fatto, la soluzione di Halmos dipendeva da quella non standard e il senso dell'articolo era di dimostrare come, in ogni caso, i concetti logici non fossero essenziali alla dimostrazione del teorema e fossero in via di principio eliminabili.

Quello di Halmos era uno dei possibili atteggiamenti che i matematici «convenzionali» hanno avuto nei confronti dell'Analisi non standard: il riconoscimento dei singoli risultati ma il sostanziale rigetto, come estraneo alla matematica, di tutto l'armamentario logico. Risposte diverse sono state date da altri matematici anche sul piano didattico. L'Analisi non standard permette un approccio diretto all'Analisi infinitesimale (il *calculus* degli anglosassoni) in cui infiniti e infinitesimi rimpiazzano le elaborate considerazioni di limiti in termini di ε e δ; non sono mancati così i manuali con tale impostazione, ad esempio *Elementary Calculus* (*Calcolo elementare*) di H.J. Keisler uscito nel 1976 e tradotto anche in ita-

liano nel 1982, e *Introduction to Non-standard Analysis* (*Introduzione all'Analisi non standard*, 1975) di W.A.J. Luxemburg e K.D. Stroyan, dove è interessante osservare che Luxemburg, a differenza degli altri due autori, è una analista di professione.

Dal nostro punto di vista ancor più significativo è stato l'impiego dei metodi non standard nella ricerca tanto pura quanto applicata. Già Robinson, non solo nel suo libro *Non-standard Analysis* (*Analisi non standard*) del 1965 ma in numerosi contributi successivi, aveva mostrato quanto vari fossero i campi d'applicazione dei nuovi metodi, dalla fisica teorica (in collaborazione con Peter J. Kelemen) all'economia matematica (con Donald J. Brown) all'Analisi funzionale (come sopra abbiamo ricordato) agli sviluppi asintotici (con A.H. Lightstone) ed infine con gli ultimi lavori dedicati alla teoria dei numeri prima della morte in collaborazione con P. Roquette. Tutto questo, inquadrando il discorso su uno sfondo storico e filosofico il cui obiettivo era di mostrare come l'impiego di metodi logici permettesse di recuperare e di rendere operative intuizioni, come quella di infinitesimo, che lo sviluppo storico aveva relegato nell'ombra ma che, sul piano teorico, risultavano ancora feconde.

Molte delle applicazioni indicate da Robinson o date dopo di lui, pongono in luce proprio questo aspetto e non deve stupire che sia stato proprio sul versante della matematica più vicina alle applicazioni – in cui l'adeguazione degli strumenti alle situazioni reali e alle intuizioni dei ricercatori è essenziale – che l'Analisi non standard ha rivelato le sue maggiori potenzialità. Questo discorso si applica in particolare a tutta l'area di problemi che riguarda la teoria della probabilità, l'analisi stocastica, lo studio dei moti browniani e più in generale la fisica matematica. In questo contesto un peso fondamentale hanno avuto i lavori di Peter A. Loeb dalla metà degli anni settanta sulla possibilità di definire misure standard a partire da misure non standard. Da queste ricerche hanno preso l'avvio indagini estremamente significative sia in campo puro che applicato e tra queste un ruolo particolare hanno i lavori di Keisler sulla *teoria dei modelli iperfiniti*, la logica della probabilità e l'Analisi stocastica.

L'idea alla base della teoria dei modelli iperfiniti è quella che abbiamo visto operante, in un certo senso, già nell'aritmetica non standard e consiste nel considerare le strutture *iperfinite* (dove un

oggetto è iperfinito quando la sua cardinalità è un naturale infinito) come strumenti per analizzare il comportamento al limite delle strutture finite. Il vantaggio che i metodi non standard offrono è che in questo modo possiamo utilizzare argomenti combinatori *finiti* in un opportuno contesto infinito. Come osserva Keisler, «nella matematica applicata spesso si approssimano fenomeni molto grandi con modelli infiniti più semplici. I modelli considerati nella letteratura sono spesso stati o successioni numerabili di strutture finite o strutture costruite a partire dai numeri reali. I modelli iperfiniti forniscono una fonte migliore di modelli infiniti che approssimano più da vicino fenomeni finiti grandi».

Con questi lavori – che proseguivano la linea iniziata da Robinson – si mostrava chiaramente come l'Analisi non standard non sia solo un mezzo per rendere più intuitive considerazioni che coinvolgono processi al limite, ma fornisca un vero e proprio contesto alternativo in cui costruire strutture che modellano in maniera più fedele le situazioni estremamente eterogenee che si presentano.

Non più solo, quindi, un potenziamento dei metodi classici, ma un ampliamento delle stesse categorie di fondo alla base della matematizzazione delle scienze empiriche. Torneremo più avanti su questo aspetto che ci sembra di estrema importanza. Quello che vogliamo osservare ora è che la necessità di rendere operante l'uso dei metodi non standard ha avuto come ovvia conseguenza la ricerca di formulazioni che si rivelassero atte a questo scopo più di quella originariamente adottata da Robinson, che si basava sul linguaggio della teoria dei tipi. L'utilizzo massiccio di questo genere di linguaggi è stato senza dubbio una delle cause non ultime della insofferenza di molti matematici nei confronti dell'Analisi non standard. Fu così che nel 1967 Robinson stesso in collaborazione con Elias Zakon presentò un nuovo approccio basato sul concetto di *superstruttura*. L'idea di fondo, che mima la struttura cumulativa dei tipi introdotta da Zermelo, è quella di considerare una successione di universi parziali $V_0 \subseteq V_1 \subseteq \dots$ dove $V_0 = \mathrm{R}$, l'insieme dei numeri reali, e per ogni n

$$V_{n+1} = V_n \cup \mathscr{P}(V_n).$$

La superstruttura su R è data dalla riunione

$$V = \bigcup_{n=0}^{\infty} V_n.$$

Gli enti studiati dall'Analisi si ritrovano tutti ad opportuni livelli di questa successione e quindi tutti appartengono alla superstruttura. Come insiemi, fra di essi è definita la relazione di appartenenza ed è facile verificare che i vari livelli non sono altro che i vari tipi considerati da Robinson nel suo approccio originario che li distingueva l'uno dall'altro anche sul piano linguistico. Nel nuovo approccio che stiamo descrivendo, questo non avviene ed il linguaggio L che considereremo non conterrà distinzioni di tipo, se non quelle dei linguaggi del primo ordine. Tra le costanti avremo nomi $\bar{c}$ per ogni elemento c della superstruttura e il simbolo ε per la relazione di appartenenza. Lo spirito dell'Analisi non standard, come sappiamo, è quello di vedere la struttura standard immersa in una soprastruttura non standard in cui appunto ritroviamo gli oggetti non standard, infiniti e infinitesimi. Accanto alla nostra superstruttura standard V consideriamo allora una nuova superstruttura non standard V^* tale che

$$V_0 \in V^*$$

e supponiamo di avere un morfismo $*$ di immersione

$$* : V \rightarrow V^*$$

tale che

per ogni $r \in V_0$, $r^* = r$
per $x, y \in V$, $x \in y$ se e solo se $x^* \in y^*$.

È immediato verificare che * sarà un monomorfismo che coinciderà con l'identità su V_0 e che ci permette quindi di vedere V come immerso nell'universo non standard V^*. Di fatto V^* non è altro che il modello non standard in cui V si immerge elementarmente e che Robinson costruiva originariamente utilizzando il teorema di compattezza per il suo linguaggio dei tipi. Quello che rimane ora

da specificare è appunto il carattere elementare di questa immersione $*$ e questo lo facciamo ricorrendo alle formule relativizzate, ossia a quelle formule in cui i quantificatori compaiono solo nella forma

$$\forall x(x \in \bar{c} \rightarrow \ldots) \text{ e } \exists y(y \in \bar{c} \wedge \ldots).$$

A questo punto la condizione che ci occorre per poter applicare le tecniche non standard è il seguente

Principio di Leibniz

$$* : V(\mathrm{R}) \rightarrow V(\mathrm{R}^*) \text{ preserva gli enunciati relativizzati}$$

dove dire che un enunciato $\mathscr{A}$ è preservato significa che se esso è vero in V sarà vero in V^* l'enunciato $\mathscr{A}^*$ che si ottiene interpretando ogni nome $\bar{c}$ come il nome di c^*.

Il principio di Leibniz è in questo contesto la traduzione del fatto che l'universo non standard V^* costituisce un'estensione elementare dell'universo standard V ed il vantaggio rispetto all'approccio originale di Robinson è che qui, ricorrendo agli enunciati limitati, possiamo fare a meno dell'ingombrante macchinario della teoria dei tipi. L'uso delle superstrutture rende particolarmente trasparente l'intergioco fra i diversi tipi di elementi che compaiono nell'universo non standard. Così possiamo distinguere gli elementi *reali* – che sono gli elementi di V, quelli *standard* – della forma a^* per qualche $a \in V$, quelli *interni* – che sono quegli enti b per cui esiste un $a \in V$ tale che $b \in a^*$ – ed infine gli enti *esterni* – che sono quelli non interni. È facile verificare che ogni oggetto standard è anche interno e che standard e reali sono tutti i numeri reali e tutti i sottoinsiemi finiti di numeri reali. Più in generale si può provare che i finiti sono tutti e soli i sottoinsiemi interni di R. R stesso, come N, pur essendo reale, è esterno, e questo riflette il fatto che passando alla superstruttura V^* avremo che N^* e R^* contengono anche naturali e reali infiniti, cosicché sono insiemi non reali pur essendo standard.

La struttura di R^* è particolarmente interessante ed è stata fatta oggetto di analisi specifiche da parte di Elias Zakon. Il primo fatto da osservare è che R^* è un'estensione *propria* di R e che quindi costituisce un campo ordinato non archimedeo. Ciò significa

che contiene infinitesimi *propri*, elementi *a* che in valore assoluto sono minori di ogni numero razionale (e quindi, per la densità, di ogni reale $\neq 0$). L'unico infinitesimo che appartiene ad R è l'infinitesimo improprio che è costituito da 0, mentre in R* gli infinitesimi coincidono con gli inversi moltiplicativi degli infiniti. Se chiamiamo *finito* ogni elemento di R* maggiorato in valore assoluto da un numero naturale, avremo che la famiglia degli elementi finiti in R costituisce un sottoanello *locale*, nel senso che in esso esiste un solo ideale massimale che è dato dagli elementi infinitesimi. È facile verificare che R sarà isomorfo, come campo ordinato, al quoziente dell'anello dei finiti modulo l'ideale degli infinitesimi, il che ci mostra tra l'altro come la struttura dei reali si possa in modo molto semplice costruire utilizzando i metodi non standard.

Passando al quoziente vengono identificati quegli elementi finiti la cui differenza è infinitesima (appartiene cioè all'ideale). In altri termini, possiamo dire che elementi della stessa classe sono infinitamente vicini e con Robinson possiamo chiamare *monade* di un elemento *a* la classe di equivalenza da esso determinata, vale a dire la classe degli elementi ad esso infinitamente vicini. Gli infinitesimi sono la monade di 0, mentre *galassia* di un elemento *a* è la famiglia degli elementi di R* che è a distanza finita da *a*. Il concetto di monade può essere esteso ad elementi arbitrari della superstruttura cosicché è possibile sviluppare una teoria generale delle monadi – come fatto da Luxemburg – offrendo un contesto particolarmente interessante per lo studio degli spazi topologici e delle nozioni connesse. Non avrebbe senso continuare ulteriormente nell'analisi della struttura di V^*; ci basti osservare che risulta chiaro come nell'impostazione sopradescritta *tutti* i concetti definiti per la struttura standard si traducano in concetti definiti nella struttura non standard. Questo ad esempio è il caso di quella nozione di iperfinito di cui sopra si è detto, che è il correlato in V^* della proprietà «essere finito» data in V.

La descrizione del concetto di superstruttura e di $*$-immersione è generale nel senso che in contesti specifici altre proprietà possono essere postulate, in particolare proprietà di saturazione, e questo evidenzia il fatto che nell'Analisi non standard non è scelto univocamente il tipo di universo non standard cui si fa riferimento, ma esso può cambiare a seconda del genere di problemi di cui ci si occupa: l'importante è che vengano sempre soddisfatte le con-

dizioni minimali poste sopra. Legato a questo è il problema del come costruire *-immersioni con particolari proprietà. La tecnica preferita da Robinson era il ricorso alla compattezza, un altro metodo sfruttato in varie maniere da Luxemburg è quello degli ultraprodotti in cui le proprietà di saturazione dell'estensione non standard sono strettamente legate alle proprietà di completezza dell'ultrafiltro modulo con il quale si ottiene l'ultraprodotto.

Uso della teoria dei tipi da una parte e passaggio alle superstrutture dall'altra non sono però gli unici modi con cui si è tentato di fornire un contesto maneggevole ove inquadrare la pratica non standard. Già dalla fine degli anni sessanta erano stati proposti metodi sintattici, con lo scopo tra l'altro di analizzare entro che limiti, una volta opportunamente assiomatizzata, l'Analisi non standard costituisse un'estensione conservativa di quella standard. In tempi più recenti Edward Nelson (uno dei più convinti sostenitori dell'importanza dei metodi non standard nella matematica e nella fisica matematica) ha proposto una sua teoria non standard degli insiemi, mentre Petr Vopěnka, partendo dagli sviluppi della teoria dei semisets di cui abbiamo già parlato, è andato proponendo (come già ricordato) una *teoria alternativa degli insiemi* ($\mathfrak{AST}$) come contesto per la pratica non standard. Discorso diverso andrebbe fatto per gli sviluppi estremamente significativi che l'Analisi non standard ha avuto nel contesto della logica categoriale ove l'attenzione è sull'aspetto geometrico degli infinitesimi.

L'esempio dell'Analisi non standard mostra in maniera molto chiara come la teoria dei modelli non sia un fatto interno della matematica che si limita ad algebra od Analisi ma – fedele al proprio nome – possa in alcuni casi offrire strumenti per «modellare» situazioni che non nascono dalla sola matematica ma da altri campi, economia, fisica, ecc. È rifacendosi a questo fatto che nel suo intervento al simposio in onore di Tarski, tenutosi nel 1971, Chen Chung Chang si chiedeva: «I nostri modelli sono sufficienti per lo studio di tutte le "scienze deduttive"?» Anche se storicamente le cose potevano stare così al tempo in cui Tarski introduceva questa teoria e possono ancora stare così per la matematica, «... ciò non è vero per il termine "modelli" come è inteso e impiegato da un gran numero di ricercatori in branche diverse quali... l'economia, la biologia, la psicologia... Tutti costoro sono alla disperata ricerca di modelli. Ognuno cerca di trovare un modello

senza che tuttavia nessuno sembri sapere che cosa esso sia; e tutti inoltre concordano sul fatto che i modelli hanno o avranno un ruolo importante nella loro scienza». Ora, aggiunge Chang, «noi specialisti della teoria dei modelli sappiamo tutti esattamente che cosa è un modello e possiamo continuare a ripeterlo a questi ricercatori. Ma i nostri modelli non li soddisfano. Ciò può avvenire o perché la loro scienza non è ancora deduttiva, o perché il loro linguaggio non è abbastanza formale, oppure (e questa è forse la ragione più cogente) perché ci sono cose intrinsecamente più complicate di quelle che noi chiamiamo modelli... Qualunque possa essere la loro nozione di modello, non si tratta di quella che noi stiamo studiando in modo così intenso».

Chang proseguiva osservando che una tendenza che sembra essersi sempre più affermata di recente nella ricerca matematica è quella di indirizzarsi verso problemi pratici, nella soluzione dei quali il matematico può risultare utile. Qualcosa di analogo, secondo Chang, doveva accadere anche per quanto riguarda gli specialisti di teoria dei modelli, e questo non nel senso che essi debbano abbandonare questa teoria così come ora la conoscono e l'hanno costruita per diventare tutti dei «modellisti applicati»; l'idea che Chang vuole suggerire è piuttosto quella che «se pensiamo di applicare la teoria dei modelli dobbiamo essere disponibili e preparati a studiare tipi di modelli certamente non ortodossi».

La teoria dei modelli iperfiniti di Keisler, con tutta la sua gamma di applicazioni, costituisce senza dubbio un esempio di ciò che Chang auspicava per quanto riguarda gli aspetti tecnici della questione, ma le parole di Chang riflettono un'esigenza profondamente avvertita anche da un punto di vista epistemologico, ponendosi dal quale l'usuale analisi logica delle teorie empiriche ha il difetto di essere, per così dire, troppo «idealizzata»: in generale infatti essa trascura tutta una serie di elementi che i vari ricercatori ritengono caratteristici o comunque essenziali nella «pratica» delle singole scienze e tali quindi da non dover essere trascurati in un'analisi teorica delle stesse.

D'altra parte, pur se con posizioni assai diverse, molti autori sono andati da tempo sostenendo l'opportunità di studiare le scienze empiriche con gli strumenti della teoria logica dei modelli; ne è testimonianza il simposio organizzato da Hans Freudenthal nel 1960 dedicato al ruolo dei modelli in matematica e nelle scienze

naturali e sociali, cui parteciparono diversi fisici e logici tra i quali E. Beth e Patrick Suppes che avanzavano il progetto di una *teoria dei modelli per le teorie empiriche*, in analogia con quella per le teorie matematiche. Suppes così sosteneva che al concetto di modello come elaborato della logica è riconducibile qualunque altra analoga concettualizzazione di scienze empiriche particolari. Per quanto riguarda la fisica invece – con molta maggior cautela – Joseph Sneed osservava che l'usuale caratterizzazione semantica coglie al più la «fisica matematica» piuttosto che la fisica *tout court* anche se così delineato il progetto di una teoria dei modelli per le scienze empiriche risultava estremamente interessante. In questo contesto una teoria fisica – esattamente come avviene per una teoria matematica – risultava determinata da una classe $\mathcal{M}$ di strutture $\mathfrak{M}$

$$(*) \qquad \mathfrak{M} = < \mathfrak{M}_0, S, q_1, ..., q_n >$$

dove $\mathfrak{M}_0$ è la parte matematica di $\mathfrak{M}$ (ad esempio un modello standard per l'Analisi), S è un insieme di oggetti fisici e $q_1, ..., q_n$ sono grandezze fisiche che formalmente possono descriversi come funzioni che a oggetti fisici (elementi di S) associano numeri reali (elementi di $\mathfrak{M}_0$).

Si è venuta così sviluppando tutta una serie di ricerche che ha come obiettivo centrale – pur partendo da prospettive diverse – l'elaborazione di una semantica per le teorie empiriche. Alcuni suoi rappresentanti sono, oltre a Sneed, Ryszard Wojcicki e Marion Przelecki in Polonia, Bas van Fraassen, P. Suppes negli Stati Uniti, M.L. Dalla Chiara e G. Toraldo di Francia in Italia che a partire dagli anni settanta hanno portato contributi interessanti all'argomento in collegamento anche col problema dei fondamenti della meccanica quantistica e della sua struttura logica, campo particolarmente complesso in cui la difficoltà centrale – come sottolineano le riserve di Sneed sopra ricordate – è data dal pericolo di ricondurre in modo troppo semplicistico le teorie empiriche e quelle formali.

Questo rischio è chiaramente denunciato da M.L. Dalla Chiara e Giuliano Toraldo di Francia in un loro articolo del 1973 dedicato appunto all'analisi logica delle teorie fisiche che, a loro parere, «vengono assai spesso trattate in un modo che assomiglia troppo da vicino all'analisi delle scienze formali, col risultato che l'aspet-

to caratteristico della fisica è semplicemente rappresentato dalla necessaria richiesta che la teoria ammetta una parziale interpretazione osservativa, che non è un modello nel senso della semantica logica. Una tale interpretazione, che dà origine al problema di separare i termini osservativi da quelli teorici,[6] non costituisce né un aspetto specifico né il più interessante delle teorie fisiche». Particolarmente significativo è il caso del concetto di *grandezza fisica* di cui le usuali caratterizzazioni come sopra descritte sembrano fare un'idealizzazione particolarmente spinta, dal momento che non tengono conto di alcuni suoi aspetti essenziali e specifici. Il problema è di «collegare l'analisi teorica con la fisica *reale*», con l'intento epistemologicamente più rilevante di confutare opinioni largamente diffuse secondo le quali «la fisica può solo formulare delle regole *strumentalmente* utili per scopi pratici ma che nulla hanno a che fare con la verità o la conoscenza»; e di mostrare, sulla base della loro analisi, che viceversa «le verità della fisica sono verità *rigorose*, nel senso della semantica logica e non si limitano ad approssimare delle verità "più profonde" o "ultime"».

Con questi temi ci avviciniamo ad una problematica molto più generale, riguardante la logica dello sviluppo delle teorie, il concetto di verosimiglianza (*truthlikeness*) ecc., problematica che negli ultimi decenni ha dato origine a dibattiti accesi e a ricerche interessanti in contesti che trascendono quello della logica per coinvolgere la filosofia della scienza e più in generale la filosofia. Al di là di considerazioni più specifiche che rimandiamo al prossimo capitolo, ci sembra comunque che tutto questo sviluppo mostri l'emergere di una tendenza unitaria, che amplia i confini dell'indagine logica abbattendo semplificazioni di tipo riduttivo.

Come ricordava recentemente Barwise citando S. Ulam, «forse è venuto il tempo di arricchire la logica formale aggiungendole altre nozioni fondamentali».

[6] Secondo gli autori questa dicotomia non è stata mai chiaramente determinata e lo stesso Carnap che l'aveva introdotta ha finito con l'annetterle solo un valore pratico; di avviso diverso Nagel che sostiene ancora la grande importanza della distinzione, che invece è semplicemente rifiutata da altri autori come Putnam. Dalla Chiara e Toraldo di Francia non solo sono del parere di Putnam (anche se per ragioni diverse) ma sostengono inoltre che «un completo rifiuto della dicotomia tra termini osservativi e termini teorici si troverà essere essenziale per una migliore comprensione del concetto di grandezza fisica» che è il concetto centrale di cui si propone una nuova analisi nel loro articolo.

G. Lolli, *Categorie, universi e principi di riflessione*, Boringhieri, Torino 1977.

Id., *Teoria assiomatica degli insiemi. Insiemi costruibili e modelli booleani*, Boringhieri, Torino 1974.

J.D. Monk, *Introduzione alla teoria degli insiemi*, Boringhieri, Torino 1972 [edizione originale 1969].

AA.VV., *Generalized Recursion Theory*, a cura di Jens Erik Fenstad e Peter G. Hinman, North Holland, Amsterdam 1974.

AA.VV., *Generalized Recursion Theory II*, a cura di J.E. Fenstand, Robin O. Gandy e Gerald E. Sacks, North Holland, Amsterdam 1978.

H.P. Barendregt, *The Lambda Calculus. Its Syntax and Semantics*, North Holland, Amsterdam 1981.

J.R. Hindley, B. Lercher e J.P. Seldin, *Introduzione alla logica combinatoria*, Boringhieri, Torino 1975 [edizione originale 1972].

S. Stenlund, *Combinators, Terms and Proof Theory*, Reidel, Dordrecht 1972.

M.L. Dalla Chiara, *Modelli sintattici e semantici delle teorie elementari*, Feltrinelli, Milano 1968.

C. Cellucci, *Teoria della dimostrazione*, Boringhieri, Torino 1978.

R. Kaye, *Models of Peano Arithmetic*, Clarendon Press, Oxford 1991.

A. Robinson, *Non-standard Analysis*, North Holland, Amsterdam 1966.

H.J. Keisler, *Foundations of Infinitesimal Calculus*, Prindle, Boston 1976.

Id., *Elementi di analisi matematica*, 2 volumi, Piccin, Padova 1982 [edizione originale 1976].

M.L. Dalla Chiara e G. Toraldo di Francia, *Le teorie fisiche. Un'analisi formale*, Boringhieri, Torino 1981.

W.S. Hatcher, *Fondamenti della matematica*, Boringhieri, Torino 1973 [edizione originale 1968].

AA.VV., *9 lezioni di logica*, Muzzio, Padova 1989.

M.L. Dalla Chiara, *La logica*, Mondadori, Milano 1974.

G. Lolli, *Introduzione alla logica formale*, Il Mulino, Bologna 1991.

J.R. Shoenfield, *Logica matematica*, Boringhieri, Torino 1970 [edizione originale 1967].

E. Mendelson, *Introduzione alla logica matematica*, Boringhieri, Torino 1972 [edizione originale 1964].

E. Bencivenga, *Il primo libro di logica*, Boringhieri, Torino 1984.

V.M. Abrusci, *Logica matematica*, Edizioni Fratelli Laterza, Bari 1992.

G. Lolli, *Capire una dimostrazione. Il ruolo della logica matematica*, Il Mulino, Bologna 1988.

Id., *La macchina e le dimostrazioni*, Il Mulino, Bologna 1987.

AA.VV., *Handbook of Mathematical Logic*, a cura di J. Barwise, North Holland, Amsterdam 1977.

AA.VV., *Handbook of Philosophical Logic*, 4 volumi, a cura di D. Gabbay e F. Günthner, Reidel, Dordrecht 1983.

R. Goldblatt, *Topoi*, North Holland, Amsterdam 1984[2].

P.T. Johnstone, *Topos Theory*, Academic Press, Londra 1977.

M. Makkai e G. Reyes, *First order categorical Logic*, Springer Verlag, Berlino 1977.

J. Lambek e P. Scott, *Introduction to higher order categorical Logic*, Cambridge University Press, Cambridge 1986.

M. Barr e C. Wells, *Toposes, Triples and Theories*, Springer Verlag, Berlino 1985.

AA.VV., *Mathematical Developments arising from Hilbert Problems*, 2 volumi, AMS, Providence 1976 (in particolare gli articoli di D. Martin, G. Kreisel, M. Davis, Y. Matijasevič e J. Robinson).

AA.VV., *Ω-Bibliography of Mathematical Logic*, a cura di Gert H. Müller, 6 volumi, Springer Verlag, Berlino 1987.

CAPITOLO OTTAVO
OSSERVAZIONI CONCLUSIVE

Giunti alla conclusione di una esposizione che certo non è stata né lineare né sempre bilanciata nella presentazione e nell'approfondimento delle diverse linee di ricerca, è opportuno fare alcune considerazioni di carattere generale. La prima osservazione da fare è che tale disparità di trattamento dei diversi argomenti dipende certo dai limiti delle competenze, degli interessi e, perché no, delle simpatie degli autori, ma ha anche un fondamento oggettivo nella natura delle cose. Lo stesso discorso si applica anche alla labilità dei confini temporali della nostra storia che solo con beneficio di inventario possiamo dire termini con i primi anni settanta. Le diverse direttive di ricerca, addirittura le diverse sottodiscipline (teoria degli insiemi, dei modelli, della dimostrazione, ecc.) della logica matematica non hanno avuto uno sviluppo uniforme perché la logica non è un blocco monolitico con ben fissati obiettivi che si possono descrivere in modo netto e definitivo; questo vale soprattutto per la ricerca dopo gli anni sessanta in cui non solo la quantità di lavori è aumentata in modo considerevole, ma anche la varietà ed i collegamenti con altre discipline, siano esse l'algebra, l'Analisi, la geometria, come pure la linguistica, la computer science, ecc. Il problema della unità della ricerca logica non è comunque un problema fittizio.

Nel 1977 usciva per i tipi della North Holland quello che ormai è un classico, un punto di riferimento obbligato per chi si avvicini oggi alla logica matematica, l'*Handbook of Mathematical Logic* (*Manuale di logica matematica*), a cura di J. Barwise. I diversi contributi ivi contenuti compongono un quadro compatto e unitario che comprende teoria degli insiemi, ricorsività, teoria dei modelli, teoria della dimostrazione, matematica costruttiva ed infine un capitolo dedicato alla logica categoriale. Il lettore, anche se percepisce differenze di prospettiva e di tematica, ricava in ogni caso un'im-

pressione globale di unità, la sensazione di trovarsi di fronte ad una disciplina ben determinata. Questa impressione, che di fatto è corretta, è nella realtà più difficile da avere per chi oggi cerchi di orientarsi nella tematica della logica matematica. Già nel 1977 c'era tutto un settore della ricerca che non compariva nell'*Handbook* ed è la sterminata serie di lavori che riguardano le logiche modali, temporali, deontiche, ecc. e le applicazioni alla linguistica. A prova della vastità di questi temi pochi anni dopo iniziavano ad uscire i quattro volumi dell'*Handbook of philosophical Logic* (*Manuale di logica filosofica*) a cura di F. Günthner e D. Gabbay. Non si trattava certo di una dimenticanza, né forse neppure di limiti dettati da esigenze editoriali, ma dell'accettazione di una dicotomia tra logica matematica e logica filosofica che era stata addirittura canonizzata da Nicholas Rescher in un suo famoso articolo del 1968, come dovuta alla necessità di far risaltare tutta una serie di altre ricerche a suo parere oscurate, in particolare fra gli anni cinquanta e sessanta, dal continuo susseguirsi di risultati di «logica matematica», che avrebbe fatto passare in secondo piano «il fenomenale scatto di crescita della logica in direzioni riferentisi a considerazioni filosofiche».

Questa suddivisione di Rescher sembrava quindi riferita non tanto alla logica matematica – che come strumento resterebbe costante in entrambi i casi – quanto all'argomento che di volta in volta viene preso in considerazione. Come abbiamo già avuto occasione di osservare in altro luogo, questa distinzione, per quanto utile e fruttuosa possa risultare per scopi pratici,[1] ci appariva allora, e ancor più ci appare oggi, comunque artificiosa: essa ripropone dicotomie di antica data e di tipo quasi sempre conservatore e reazionario, così ben tratteggiate da Frege nei due contrapposti atteggiamenti alla *metaphisica sunt non leguntur* o alla *mathematica sunt non leguntur*, che egli per primo intraprese a combattere con l'impostazione filosofica di un tipico programma tradizionalmente «matematico».

[1] Essa può avere un puro e semplice valore classificatorio convenzionale allo scopo ad esempio di «ordinare» una rassegna bibliografica, o di mettere in evidenza quelle che possono essere recenti applicazioni della logica in questo o quel campo culturale; o al limite per dare il nome a una rivista, o altro. Ma non ha comunque, in questo concordiamo pienamente con Hintikka (si veda più avanti), alcun valore né giustificazione teorici.

A noi inoltre richiama alla mente infeconde dicotomie pratico/teorico tra i fautori dell'essere e i dominatori del reale e la riproposizione della separazione fra «scienze esatte» e «filosofia» che – a nostro giudizio – appartengono proprio al modo più deleterio e sterile di vedere la filosofia. In realtà quello che crediamo debba risultare dall'esame dello sviluppo della logica che abbiamo qui tratteggiato è che proprio la logica matematica ha contribuito fortemente a colmare per certi aspetti quell'arbitraria e improduttiva frattura che in varie culture veniva (e viene) operata fra filosofia e matematica; ha cioè permesso di gettare un ponte fra due «attività», per dirla con Wittgenstein, che di diverso possono eventualmente avere metodi e problemi, ma non certo «dignità» culturale o «profondità» di discorso. In altri termini, forse un po' semplificati, riteniamo che oggi la logica sia «matematica» per tutta una serie di complesse motivazioni storiche che abbiamo cercato di illustrare e non in quanto contrapposta a «filosofica».

Già nel 1973 scriveva Hintikka: «Esiste qualcosa come la "logica filosofica"? Fondamentalmente la mia risposta è "no". Non sembra esserci dal punto di vista dell'interesse filosofico molta differenza intrinseca tra i diversi rami della logica. Buona parte del lavoro recente nelle parti più esoteriche della logica matematica ha, a mio avviso, una grande rilevanza per l'indagine filosofica. È vero che la maggior parte di questo lavoro non ha richiamato l'attenzione dei filosofi o, per lo meno, che questo è avvenuto soltanto in quei casi relativamente rari in cui i logici stessi hanno richiamato l'attenzione dei filosofi verso i loro problemi. Tuttavia, gran parte del lavoro che è stato fatto in campi come la teoria delle funzioni ricorsive, la teoria dei modelli e la metamatematica riguarda l'esplicazione e lo sviluppo di concetti e problemi concettuali che sono del massimo interesse e della massima rilevanza agli scopi del filosofo».[2]

[2] Alle esemplificazioni di Hintikka si può aggiungere senza difficoltà tutta la problematica relativa alla teoria delle categorie nel suo riflesso sulla logica e sulla concezione stessa della matematica, come pure quella relativa all'analisi logica delle scienze empiriche. Mette conto comunque di riportare la conclusione del discorso di Hintikka: «Questa mancanza di interesse da parte dei filosofi verso alcune delle parti della logica più genuinamente filosofiche è naturalmente in parte inevitabile. Essa è dovuta alla difficoltà di approfondire un campo che si è sviluppato con grande rapidità e che in molti casi richiede un'alta competenza tecnica. Quello che è veramente sconcertante non è tanto l'ignoranza che si ritro-

Per parte nostra non siamo minimamente dell'opinione che l'intervento dell'apparato logico-matematico si esaurisca nel conferire «rigore» alla formulazione di posizioni e problemi in prima istanza intuitivi e non accettiamo in alcun modo la tesi talora accreditata fra i filosofi, ma non solo fra essi, secondo la quale compito della logica (e della scienza in generale) è quello di fare un discorso *rigoroso*, mentre è caratteristica della filosofia di fare un discorso *veritativo*. Il compito della logica matematica, a nostro parere, non si limita affatto a quello di conferire rigore; essa piuttosto offre concettualizzazioni e metodi che sono sì rigorosi, ma diventano operativi e fecondi solo una volta che diano origine a nuove formulazioni e interpretazioni intuitive cui devono fornire articolazioni, ma mai solo una veste esterna. Detto in termini leggermente differenti: siamo convinti che l'analisi logica *non* esaurisca in generale la comprensione estensiva del contenuto e delle potenzialità di un problema; ma siamo altrettanto convinti che essa sia *inderogabile* in quanto non di rado conduce a nuove generalizzazioni, assai precise sì, ma non meno feconde, proprio perché mette in luce momenti contenutistici fondamentali.

Con ciò non vogliamo naturalmente dire che la moderna logica matematica sia divenuta una specie di toccasana universale, con cui affrontare e risolvere «tutto». Intendiamo invece proprio il contrario: che è perfettamente inutile sognare la costituzione di strane «logiche filosofiche» con le quali si penserebbe appunto di poter risolvere ogni problema; e che ogni periodo storico ha i suoi strumenti d'indagine *specifici* che come tali vanno assunti proprio per poterli superare, migliorare e affinare continuamente, e verso i quali è ormai perlomeno assurdo avere la repellenza della «tecnicità». A nostro parere tutta questa problematica è una ulteriore dimostrazione della rilevanza filosofica che la logica matematica – o se si preferisce la trattazione matematica della logica – intesa come un *modo oggi imprescindibile di fare logica* (di qualunque tipo poi questa sia) ha in quanto tale, senza cioè che ci sia bisogno di scomodare una pretesa «logica filosofica» che non sia una pura etichetta di comodo.

va nelle discussioni di alcuni filosofi su risultati ben noti come quelli di Gödel, quanto la loro incapacità (o riluttanza) di seguire Socrate e ammettere la vastità della loro ignoranza.»

Siamo dell'avviso che un modo per uscire da questa dicotomia astratta sia di allargare il raggio dell'analisi e tenere esplicitamente conto dei legami che la logica ha storicamente intrattenuto e storicamente intrattiene con altre pratiche scientifiche, e non solo scientifiche, e vedere in che misura essa è in grado di rispondere alle esigenze che queste pratiche impongono. In altri termini si tratta di riconoscere che l'indagine logica «deve» avere a che fare con i fondamenti delle pratiche teoriche e questo «deve» nasce dall'opportunità, oggi sempre più avvertita ma ben viva anche nel passato, di un inquadramento generale del lavoro teorico, prendendo atto, nel contempo, che non sempre questo storicamente è avvenuto. Ci sono state esclusioni, pratiche teoriche privilegiate, sovrapposizioni ideologiche, il cui studio a nostro parere costituisce uno dei compiti dello storico della logica che non voglia limitarsi a registrare il sorgere e lo svilupparsi dei concetti al centro dell'indagine attuale; proprio però questa accettazione di un concetto così ampio, così «filosofico», di logica, pone in primo piano il rapporto concreto che storicamente si stabilisce fra le varie pratiche teoriche e la riflessione logica. Quest'ultima si realizza solo nella misura in cui è in grado di entrare in uno scambio effettivo con queste pratiche, e il suo valore filosofico sta nella attuazione concreta – materiale e non ideale – del suo progetto: se la logica infatti esiste come progetto, come particolare modo di affrontare il lavoro teorico, non ha contenuti che non siano simultaneamente contenuti di altre scienze o comunque attività umane: *logicus purus asinus est*, ci avverte il detto rinascimentale.

Torneremo più avanti su questa contrapposizione. Quello che vogliamo sottolineare è che la sua problematicità salta all'occhio di chiunque confronti non l'indice, ma i *contenuti* dei due *Handbook*: vi troverà sovrapposizioni e collegamenti che non sono casuali ma al contrario sono oggi al centro di ricerche estremamente interessanti.

Ciò non toglie che il problema della unità della ricerca logica rimanga e sia ben reale, al di là di contrapposizioni di tipo ideologico, come quella fra logica matematica e logica filosofica o quella assai recente fra logica matematica e logica computazionale. Quell'immagine di unità che l'*Handbook* di Barwise poteva dare, si scontra oggi con una ricerca che in questi ultimi quindici anni è andata sempre più diramandosi, instaurando contatti nuovi con discipline collaterali, specializzandosi, cosicché è spesso impossi-

bile in concreto, al di là di ogni prevenzione *a priori*, riuscire ad avere una visione adeguata di tutto quello che avviene.

Questo è uno dei motivi – a nostro parere oggettivamente fondato – che ci ha convinti ad interrompere la nostra esposizione agli inizi degli anni settanta, sacrificando temi che pure nell'edizione precedente avevamo toccato (quali ad esempio la logica categoriale) che ha senso affrontare oggi solo in modo più dettagliato e preciso. Non ci sembra che abbia senso fare «storia» di qualcosa che è in continua evoluzione senza nello stesso tempo discriminare in modo arbitrario, dati i limiti oggettivi delle nostre competenze e la massa veramente notevole di lavori assai interessanti e disparati che devono essere considerati.

Una misura dell'ampliarsi e del differenziarsi del volume della ricerca logica in questi ultimi vent'anni è data banalmente dal numero di nuove collane di testi di logica; pensiamo ad esempio alle *Oxford Logic Guides* diretta da Scott a partire dal 1977, o alla serie *Perspectives in Mathematical Logic* pubblicata da Springer a partire dalla metà degli anni settanta o ancora alle diverse collane di logica della casa editrice Bibliopolis dedicate rispettivamente alla storia della logica, alla teoria della dimostrazione e alla logica filosofica e alla linguistica formale, cui si è affiancata una ulteriore collana di *Testi per lo studio della logica matematica*, tanto per limitarci ad *alcuni* esempi significativi. Discorso analogo si può fare per le riviste, che hanno ampliato la loro diffusione e sono aumentate di numero, pensiamo ad esempio a *History and Philosophy of Logic* diretta da Ivor Grattan Guinness o alla più recente *Modern Logic*, diretta da Irving Anellis.

Tutti questi fatti (per tacere delle alterne vicende riguardanti lo status accademico della disciplina nei diversi paesi) testimoniano di una situazione generale che assimila la logica matematica ad altre discipline di carattere scientifico di più consolidata tradizione. Il consolidarsi come disciplina sul piano accademico e manualistico, il suo affacciarsi (anche da noi) nell'insegnamento preuniversitario, ha coinciso a volte con una perdita di unità sul piano teorico, o meglio con una diversificazione che spesso non è completamente dominabile. Malgrado la situazione sia così complessa, ci sembra possibile individuare alcune direttive, alcune linee di tendenza nelle indagini di quest'ultimo ventennio, su cui vogliamo brevemente soffermarci.

La prima osservazione da fare è che al di là di ogni dubbio due sono gli eventi di maggiore novità maturati dagli anni settanta in poi che sono comparsi nell'orizzonte della logica matematica: da una parte il sempre maggiore intervento della teoria delle categorie nella quasi totalità delle discipline logico-matematiche e il suo presentarsi con un ruolo esplicitamente fondazionale in contrapposizione alla teoria degli insiemi e dall'altra l'interscambio sempre più accentuato della ricerca logica con l'informatica. Con questo non vogliamo naturalmente affermare che i risultati più rilevanti di questi ultimi anni siano stati ottenuti *solo* in questi settori; vogliamo solo dire che essi rappresentano lo stacco più significativo e a nostro parere promettente rispetto alla ricerca logico-matematica tradizionale.

Nel caso della teoria delle categorie questo non deve stupire perché nell'ultimo cinquantennio, da quando cioè nel 1945 Eilenberg e MacLane gettarono le basi della teoria, essa si è andata sempre più rivelando come un linguaggio universale, una sorta di logica interna di alcuni settori di punta della ricerca matematica: la topologia algebrica, la geometria algebrica (e qui sono stati essenziali i lavori di Grothendieck da cui ha preso le mosse Lawvere per la creazione della logica categoriale) l'algebra omologica, ecc. sino a tempi recenti con le applicazioni nel campo dell'Analisi, della combinatorica, ecc. Il confronto della logica matematica con la teoria delle categorie non è quindi riconducibile ad una pura e semplice invasione della logica da parte dei metodi categoriali. Questa c'è indubbiamente stata e ha coinvolto tutti i rami della ricerca, dalla teoria degli insiemi alla ricorsività, alla teoria della dimostrazione, alla teoria dei modelli e alla matematica non standard per finire con l'analisi categoriale della logica stessa.

Al di là dell'interesse – veramente notevole – dei risultati raggiunti in questo campo specifico e che coinvolgono in modo massiccio anche l'informatica teorica, quello che ci sembra importante è che in un certo senso il contatto con la teoria delle categorie rappresenta per la logica matematica un modo di riavvicinarsi in maniera diretta alla matematica come oggi viene *concretamente* sviluppata. Come ammesso da uno dei protagonisti della ricerca logica come Kreisel alla conferenza in memoria di Brouwer del 1981, esiste di fatto una discrepanza tra «la matematica reale» e l'immagine di essa che compare nel lavoro dei logici. Questa non è

una sensazione isolata del solo Kreisel ma riflette un senso di disagio provato anche da altri personaggi chiave della ricerca contemporanea e a cui ciascuno ha dato una risposta diversa, basata su specifiche analisi del disagio stesso.

La proposta della logica categoriale parte essenzialmente dall'idea che la logica matematica come finora è stata praticata si è sviluppata in stretta solidarietà con la matematica su base insiemistica e come tale trova i suoi ostacoli e si rivela inadeguata proprio come in altri settori più interni alla matematica la tradizione insiemistica ha dovuto cedere il passo a nuovi metodi basati sulla teoria delle categorie. Il successo che l'uso delle categorie in indagini tradizionalmente logiche quali la matematica costruttiva e la teoria dei modelli ha conosciuto, sembrano essere la prova che effettivamente, ponendosi in un orizzonte categoriale e abbandonando le limitazioni di tipo insiemistico – sostituendo ad esempio topos arbitrari agli universi – la logica possa trovare sviluppi di tipo nuovo. È soprattutto in questo senso a nostro parere che acquista plausibilità l'idea di un ruolo fondazionale della teoria delle categorie, non come sostituzione del discorso insiemistico, ma come approfondimento del suo significato e allargamento a situazioni in cui l'analisi insiemistica risulta artificiosa.

Questo tipo di sviluppi dell'indagine però sono solo un aspetto di un problema più generale: quello della necessità di una maggiore adeguazione dell'analisi logica alla matematica nella sua realtà specifica al di fuori di formulazioni di tipo preconcetto. In questa direzione si muovono anche altre ricerche, non necessariamente collegate alla logica categoriale: pensiamo alla *model-theoretic algebra* che in questi ultimi anni ha avuto notevoli sviluppi, alle applicazioni alla geometria reale (la geometria sul campo dei reali, in cui i metodi di teoria dei modelli hanno svolto un ruolo veramente importante) o agli sviluppi dell'Analisi non standard, in particolare ai recenti lavori di Keisler sull'Analisi stocastica. Questo per citare solo alcuni esempi di cui abbiamo parlato nella nostra storia.

Ci sembra che anche l'altro elemento nuovo, costituito dall'intergioco con l'informatica, presenti questo aspetto di ricerca di un contatto più diretto con le situazioni reali, che raffini modellizzazioni troppo generali, e quindi, in qualche senso, riduttive. Come usare le regole nel «meccanizzare» le dimostrazioni e nel definire procedure di calcolo è stato sicuramente, almeno dagli anni tren-

ta, uno dei temi centrali della ricerca logica. Nella nostra storia abbiamo visto alcuni dei modelli che sottostanno alle diverse scelte operate nel formulare calcoli deduttivi e nel definire nozioni di computabilità. Non si può però sostenere che l'obiettivo di tali scelte o definizioni fosse quello di ottenere strumenti realmente «implementabili». Per Hilbert, per cui pure era centrale il concetto di teoria deduttiva, quello che interessava era la *dimostrabilità*, non la struttura reale delle *dimostrazioni*, e anche per Gentzen, che pure poneva le dimostrazioni al centro del suo interesse, l'obiettivo non era quello di conciliare il massimo di perspicuità col massimo di informatività: le dimostrazioni senza tagli di Gentzen non sono affatto le più economiche e certamente nemmeno le più perspicue.

C'era sicuramente il problema della difficoltà *tecnica* di elaborare strumenti deduttivi e di calcolo *efficienti* e c'era anche una mancanza di adeguata sperimentazione con i problemi di calcolo (esperienza che verrà invece resa possibile dall'avvento e dall'impiego dei calcolatori), ma c'era soprattutto una prospettiva teorica per cui quello che contava era isolare in modo corretto e matematicamente perspicuo concetti di significato *epistemologico* rilevante. La lunghezza delle dimostrazioni o la complessità dei calcoli sono incidenti che non riguardavano il problema essenziale e anche quando Gödel nel 1936 affrontò il problema della lunghezza delle dimostrazioni, lo fece solo per analizzare da una nuova prospettiva questioni riguardanti la dimostrabilità in sé e per sé.

Dal punto di vista dell'informatica, naturalmente, l'enfasi è diversa, ed è concentrata sull'implementazione, non solo dal punto di vista pratico ma anche nei suoi aspetti teorici. La complessità delle computazioni non è più un incidente, ma un tema di primaria importanza. Il fatto va ribadito: non si tratta esclusivamente di una questione pratica ma di una dimensione delle computazioni (ma lo stesso vale per le dimostrazioni) che l'analisi logica della computabilità in termini di teoria delle funzioni ricorsive aveva sostanzialmente trascurato, anche se fin dagli anni cinquanta studiosi come Hao Wang avevano attirato l'attenzione sull'importanza dei limiti umani, reali, posti ai processi di computazioni. In termini espliciti ci pare che quando si parla della matematica, della sua natura e dei suoi metodi, non si deve dimenticare questa dimensione di «operabilità». Il fatto che dei calcoli siano troppo

complessi, che i dettagli formali diventino sempre più minuti, in parole povere la «difficoltà» materiale di orizzontarsi in un determinato contesto non sono questioni irrilevanti o messe in luce solo con l'avvento dell'informatica «pratica». La considerazione della facilità o meno di maneggiare certi strumenti teorici ha sempre avuto un ruolo importante nella scelta delle articolazioni concettuali del matematico in quanto nessun matematico – diversamente da un filosofo – può dimenticarsi che la matematica è un'attività fatta da uomini. Quello che l'informatica ha portato di nuovo è da un lato una grande possibilità di sperimentazione e dall'altro nuove coordinate rispetto a cui misurare la «dominabilità» di problemi e procedure.

La stessa nozione di dimostrazione è coinvolta in questioni di questo tipo: entro che limiti ha senso considerare dimostrato un teorema come quello dei quattro colori (che afferma che *ogni* mappa può essere colorata distinguendo regioni adiacenti con solo quattro colori) la cui dimostrazione consta di una parte essenziale dovuta all'impiego di un calcolatore? Dimostrazioni del genere non sono sporadiche, ma compaiono spesso, ad esempio nella teoria dei numeri e nella matematica combinatoria. Per converso, di fronte alle quindicimila pagine che contengono la soluzione completa del problema della classificazione dei gruppi semplici finiti, in che senso possiamo parlare di risultato *conosciuto*? Conosciuto da chi? Questo ci dà la misura di come, volenti o nolenti, una dimensione antropologica nel concetto di dimostrazione esiste e questo non significa in alcun modo un «cedimento» a forme di psicologismo o di puro efficientismo. Significa a nostro parere che ogni classificazione dell'evidenza, della plausibilità dei metodi cui l'indagine sui fondamenti ci ha abituato, non può prescindere dal considerare accanto all'adeguatezza, ai dati intuitivi e alle situazioni reali, la complessità dell'uso della strumentazione concettuale adottata. Nella matematica è sempre estremamente difficile distinguere ciò che è dato da ciò che è strumentale, quello a cui dobbiamo adeguarci da quello con cui ci adeguiamo. È questo uno dei problemi filosofici che la matematica pone e che spesso il filosofo puro non avverte nemmeno.

Lo sviluppo della teoria della complessità in questi ultimi anni mostra abbondantemente come considerazioni di questo tipo possano portare a un reale approfondimento dell'analisi della compu-

tazione ponendone in luce aspetti del tutto nuovi. Da un lato nell'ambito del classicamente decidibile si comincia a distinguere una decidibilità polinomiale (vale a dire ove le risorse temporali o di altro tipo necessarie alla decisione crescono in funzione polinomiale dei dati) da una decidibilità esponenziale (nella quale tali risorse variano esponenzialmente con i dati) e in questo contesto si pongono nuovi e importanti problemi teorici come l'ormai classico $N \stackrel{?}{=} NP$ riguardante la coincidenza o meno di problemi risolubili deterministicamente o non. D'altro lato il concetto di complessità si presta ad analisi estremamente interessanti in termini di teoria dell'informazione e le ricerche di G. Chaitin hanno dimostrato come essa si possa vedere come una misura della casualità così che teoremi limitativi come quelli di Gödel vengono sostanzialmente ad affermare che «nessun numero non può dimostrarsi casuale a meno che la complessità del numero non sia minore di quella del sistema stesso» (considerato come un programma). Da questo scende che per dimostrare che un programma è minimale per ottenere una determinata successione di cifre (un certo output) occorre un sistema di complessità maggiore. Da un'altra prospettiva queste ricerche mostrano come in un certo modo, portando in primo piano la nozione di complessità, si possa tentare di definire nozioni di contenuto informativo per teorie che non siano banalizzabili come quelle tentate negli anni cinquanta da Carnap e da Popper.

Parlando di risorse possiamo accentuare addirittura in direzione fisica il senso del problema e cercare di affrontare la questione della meccanizzabilità o meno di certi processi di computo, tentando di misurare tempi, struttura di calcolo e struttura *fisica* degli strumenti che devono eseguirli. Esistono ricerche e risultati interessanti in questo senso, ad esempio di Mundici. Ma il discorso si può estendere ancora in un'altra direzione che coinvolge problematiche logiche più tradizionali. In anni recenti Girard – che abbiamo avuto occasione di nominare più volte nella nostra storia – è andato sviluppando un sistema di logica lineare che si presenta come più basilare rispetto alla logica classica e a quella intuizionista, nella quale in primo piano è messo il concetto di *risorsa* concentrando l'attenzione sulle «proprietà che si utilizzano quando sono necessarie», sulla rilevanza delle premesse per l'ottenimento della conclusione. Compaiono connettivi diversi rispetto a quelli

cui siamo abituati, la congiunzione ad esempio dà origine a due connettivi, il «per» e il «con» come pure si introduce una negazione «nil» che comporta una sorta di scambio «azione/reazione» (ad esempio il pagare in rapporto al comprare, il leggere allo scrivere, ecc.). Per questa logica esistono tanto formulazioni in termini di sequenze quanto analisi semantiche che pongono in luce legami estremamente interessanti col concetto di *entailment*, o con la ortologica di Goldblatt che compare nell'analisi logica della meccanica quantistica. Anche qui nei confronti delle ricerche tradizionali sull'*entailment* è interessante il cambio di prospettiva: il confronto con la logica classica non avviene in termini puramente contrappositivi bensì in stretto legame con problematiche nate dall'informatica, ad esempio col Prolog.

Anche per quanto riguarda il Prolog, e più in generale la programmazione logica, gli ultimi vent'anni hanno mostrato uno strano miscuglio tra elementi di novità e recupero di temi e risultati tradizionali della logica. Non a caso un ruolo fondamentale in questi linguaggi hanno quelle formule di Horn che erano emerse negli anni cinquanta nello studio delle formule preservate dai prodotti diretti; ma il fatto più importante sicuramente è il recupero dei metodi e del teorema di Herbrand lungo un'ottica che era stata già messa in luce da A.J. Robinson negli anni sessanta col suo principio di risoluzione.

Non avrebbe senso proseguire oltre nell'elencazione dei punti di contatto che sempre più vanno emergendo tra ricerca logica e informatica (e non solo questa se si pensa ad esempio agli sviluppi recenti della computer algebra in cui si incontrano informatica, teoria dei modelli e algebra, o della dimostrazione automatica in geometria che ha riportato a nuova vita le indagini tarskiane sui fondamenti della geometria). Vogliamo solo ricordare ancora un tema che nella sua generalità mostra il nesso profondo che si è realizzato in tempi recenti fra informatica, logica e teoria delle categorie: alludiamo allo studio dei funzionali di tipo superiore e allo studio dei principi del punto fisso e in generale alla semantica denotazionale che hanno una grande importanza nella questione della correttezza dei programmi ed in cui si riprendono le ricerche di D. Scott e J. Strachey sui modelli del λ-calcolo utilizzando l'armamentario concettuale della teoria delle categorie.

Tornando al nostro discorso generale, tanto i collegamenti con

la teoria delle categorie che quelli con l'informatica, abbiamo detto, illustrano una tendenza comune che è quella di recuperare un approccio più diretto alle diverse pratiche scientifiche cui la logica si rivolge, siano esse la matematica o la computer science o anche la linguistica che a sua volta, ad un livello diverso di integrazione, ha realizzato con la logica matematica dei rapporti stabili, che hanno portato allo svilupparsi di svariati tentativi di semantica formale dei linguaggi naturali.

Anche in questo caso assistiamo ad una sorta di ripensamento delle ipotesi di fondo che la logica matematica ha implicitamente assunto dalle sue origini nei confronti del linguaggio. Non si tratta solo di un allargamento del raggio d'applicazione dei metodi logici, soprattutto nella semantica dei linguaggi naturali, ma di un mutamento di prospettiva che riguarda la natura stessa di concetti basilari quali quelli di significato, denotazione, verità. Dagli anni sessanta si era sviluppato tutto un filone di ricerche che sostanzialmente trasferiva i metodi della semantica intensionale ai linguaggi naturali lungo linee che sono in ultima analisi quelle di Tarski. Oggi l'atteggiamento è sensibilmente mutato e compaiono semantiche che si basano su nozioni diverse, quali ad esempio la *situation semantics* di Barwise e Perry o la *data semantics* di F. Veldman. La prospettiva comune è quella di un maggiore rispetto per tutte quelle condizioni che riguardano il fatto linguistico e che tradizionalmente l'indagine logica escludeva: il contesto, i rapporti pragmatici fra situazioni, parlante, ascoltatore, ecc.

Nel loro libro *Situations and attitudes* (*Situazioni e atteggiamenti*) del 1983 Barwise e Perry presentano un programma di indagine semantica che più che a Frege sembra rifarsi a Peirce e alla sua concezione multidimensionale del rapporto significato/significante. Questo può stupire se si pensa che il Barwise in questione è lo stesso di cui abbiamo avuto occasione di parlare nel contesto di tutt'altro tipo di ricerche, ma testimonia il sorgere di un atteggiamento più comprensivo nei riguardi dei fatti linguistici e più critico rispetto alle assunzioni ormai tradizionali sulla natura del significato, gli atti mentali, i comportamenti, ecc. In questo i lavori di Barwise e Perry si incontrano con quelle ricerche condotte da Dennett, Stalnaker ed altri, che partendo da una visione più duttile della logica e dei suoi rapporti con l'informatica tentano l'edificazione di una più adeguata filosofia della mente.

Un fatto avrà sicuramente colpito il lettore ed è il silenzio che sinora abbiamo mantenuto su un aspetto da sempre centrale della ricerca logica, ossia la filosofia della matematica e più specificamente l'indagine sui fondamenti. A nostro parere non ci si può nascondere che questi contatti più stretti fra logica e matematica da una parte e logica e informatica dall'altra testimonino un atteggiamento per lo meno problematico rispetto all'indagine sui fondamenti. Non è un caso, ad esempio, che una rivista importante come gli «Annals of Mathematical Logic» da alcuni anni abbia cambiato il nome in «Annals of pure and applied Logic» dando ampio spazio all'informatica, come pure che un'altra rivista, l'«Archiv für mathematische Logik und Grundlagenforschung», si sia trasformata semplicemente in «Archiv für mathematische Logik». Sembra così che in un certo modo la logica matematica abbia consumato una rottura definitiva con tutto ciò che ha a che fare con la filosofia e specificamente con la filosofia della matematica: da una parte la logica matematica pura e semplice con i suoi rapporti privilegiati con l'informatica, l'algebra, ecc., dall'altra la logica filosofica che diventa studio della semantica dei linguaggi naturali o comunque come tale tende a presentarsi.

Diversi fattori hanno contribuito a questo. Uno sicuramente è che la maggior concretezza dei legami con le diverse problematiche che la logica è andata realizzando ha portato di per sé a una frantumazione o comunque a un drastico ridimensionamento delle classiche formulazioni programmatiche che nel tempo hanno progressivamente perduto la loro presa sulla scienza reale. Questo vale in particolar modo per la matematica, cui la veste logicista, formalista o anche intuizionista non ha tardato a mostrarsi troppo stretta. Da questo punto di vista non ci sembra ci sia nulla da rimpiangere.

La situazione cambia se prendiamo in esame altri fattori di questa evoluzione. Sicuramente l'abbandono del gemellaggio con l'indagine sui fondamenti ha avuto altre cause più contingenti, non ultima il desiderio di uscire dall'isolamento all'interno della comunità matematica, sacrificando l'aggancio con un tipo di problematica che non ha mai incontrato particolare simpatia da parte dei matematici. Da questo punto di vista l'incontro con l'informatica ha spesso avuto un doppio ruolo: quello di garantire credibilità scientifica al logico, che finalmente si occupava di «cose serie»

e dall'altra quello di permettere al matematico stesso di avvicinarsi all'informatica (forse, meglio, alle sue fonti di finanziamento) passando per la logica. Dall'altro lato, sul versante degli informatici, la logica è stata spesso il cavallo di Troia per darsi una credibilità matematica, lettera di presentazione che ormai sempre più tende a rivelarsi come non necessaria in quanto gli informatici sono in grado essi stessi di farsi la propria logica. Minore è stato il numero dei logici che decisamente sono passati armi e bagagli alla matematica in senso tradizionale portandovi il loro patrimonio logico che per il matematico conserva pur sempre un sapore sospetto, irrimediabilmente filosofico.

Quello che ci sembra paradossale in questa situazione è che – a nostro parere – mai come oggi nel vivo della matematica, dell'informatica e dello studio del linguaggio si presentino pressanti problemi di tipo filosofico e fondazionale, in qualunque senso si voglia intendere quest'ultimo termine. Non è un caso che proprio da matematici come MacLane, Lawvere o Giancarlo Rota, per fare solo alcuni esempi, che pur da prospettive molto diverse hanno condotto una lotta contro la tradizionale indagine sui fondamenti, sia venuto più volte un forte appello ad indagare la filosofia della matematica ponendo in luce questioni e temi rilevanti.

Questo singolare tiro incrociato in cui la filosofia viene ora richiesta ora rigettata porta in primo piano quello che secondo noi è il problema preliminare da affrontare: dove dobbiamo cercare la filosofia della matematica, nelle dichiarazioni programmatiche o nei fatti che la matematica presenta? La contrapposizione non è retorica. Se consideriamo i grandi programmi fondazionali presenti agli inizi del secolo e vediamo quello che di essi oggi è rimasto vivo e operante, non possiamo sottrarci all'impressione che spesso i drastici contrasti tra schieramenti si sono rivelati fittizi anche se, va sottolineato, sono forse stati necessari per porre in luce contrapposizioni reali. Per considerare l'esempio più estremo: l'intuizionismo, che si presentava come un'alternativa radicale al modo classico di fare matematica, si è rivelato negli ultimi anni la forma naturale di matematica possibile in un topos (una particolare categoria ricca come l'universo degli insiemi) i cui oggetti sono fasci o prefasci, cioè enti propri della matematica classica, e come tale svolge un grande ruolo entro quest'ultima.

Questo non significa che l'esigenza costruttivistica non abbia

senso, significa solo che si presentano punti di vista in cui queste contrapposizioni si ricompongono in un'unità superiore. Vale naturalmente anche il discorso contrario: concezioni che sembravano solidali rivelano contrapposizioni più profonde e reali. È il caso delle alternative che lo sviluppo della matematica costruttiva è andata sempre più mostrando e in un senso anche quello della concezione insiemistica. L'analisi sistematica dei vari possibili universi di insiemi non può non ingenerare il sospetto che la nozione di tali universi non sia univoca e che ci siano tante alternative possibili, adeguate ciascuna a modellare diversi tipi di situazioni.

In altre parole, ci sembra di poter concludere, più che di un rigetto di posizioni filosofiche siamo in presenza di uno spostamento dell'ambito in cui cercare *materiali* per una filosofia della matematica. Tutto sommato i classici programmi fondazionali accettavano come punti nodali nel dibattito sulla natura della matematica i temi tradizionali della natura degli enti matematici e delle procedure dimostrative. Su questi temi avveniva la separazione, ma è proprio qui che oggi è difficile ritrovare motivazioni plausibili e significative. Malgrado l'opinione di numerosi filosofi (pensiamo ai recenti dibattiti sul realismo in matematica in cui sono intervenuti filosofi come Hilary Putnam, Hartley Field, Jean Bigelow, Penelope Maddy) è difficile vedere nel problema dell'esistenza degli enti matematici una chiave per capire il senso delle diverse concettualizzazioni che le indagini matematiche rivelano.

Paradossalmente ci sembra che la causa principale di questa incapacità di coniugare ricerca filosofica e matematica stia nella convinzione che debba necessariamente esistere una disciplina a se stante, la filosofia della matematica, il cui scopo è di spiegare al colto e all'inclita che cos'è la matematica, come funziona, di che oggetti parla, ecc. come se per la matematica, e per molte altre discipline, la politica, l'arte, ecc., dovesse sempre esserci la necessità di una replica spettrale: «la filosofia della...». Messa la questione in questi termini, tanto il filosofo che il matematico tentano, spesso con protervia, di far prevalere la propria prospettiva: il matematico affermando che sa ben lui che cos'è la matematica per il semplice fatto che la fa e che quindi quelle del filosofo sono riflessioni inutili e addirittura fuorvianti, mentre il filosofo ribadisce che il matematico è un puro tecnico incapace di delibare le squisitezze filosofiche racchiuse da un guscio specialistico.

Non siamo del tutto convinti che ci sia una ragione per cui debba esistere una filosofia della matematica come *disciplina*, crediamo piuttosto che ciascuno – nel nostro caso il filosofo – *all'interno della sua problematica*, ha senso che si ponga domande sulla matematica. Non si tratta di chiarire che cosa è la matematica, perché questo probabilmente è un interrogativo senza risposta univoca, ma di accettare con umiltà che *prima* occorre aver ben chiaro il tipo di problema di cui ci si intende occupare, *poi* ha senso rivolgersi alla matematica per vedere se e come essa offra materiali utili per la risoluzione del problema dato. Per fare solo un esempio abbastanza casuale ma significativo: se un filosofo è interessato all'uso dei simboli, al ruolo che hanno nel pensiero umano, sicuramente avrà interesse a volgersi alla matematica che presenta esempi significativi di questo uso. In questa indagine dovrà essere lui a dettare le direzioni di ricerca perché suo sarà il problema e certamente non si tratterà di determinare la natura della matematica, ma di vedere come l'*esempio* della matematica ha rilevanza in una problematica filosofica già articolata.

Quello che succede spesso invece è che – nella convinzione, forse improbabile, che debba esistere una filosofia della matematica in sé e per sé – il filosofo spesso si illude che debba essere la matematica stessa a parlare da sé, ad enucleare naturalmente i problemi, col risultato, non infrequente, di prendere grossi abbagli e di dare dignità filosofica a problemi matematici non afferrati nella loro reale natura, o – se deluso – di rigettare la matematica come pura tecnica, priva di rilevanza speculativa. Quello che è grave in questo, a nostro avviso, non è tanto l'immagine falsa che può risultare della matematica, ma la perdita di una autentica dimensione filosofica, una sorta di sudditanza del filosofo rispetto al matematico che non ha ragion d'essere e che non è utile per nessuno. È una situazione che ci sembra di ritrovare spesso anche nei cultori della cosiddetta logica filosofica, in cui l'ansia di fare il matematico, di dare veste matematica alla propria indagine, porta a volte ad una totale perdita di rilevanza filosofica (accompagnata spesso da un ridicolo quanto inutile spiegamento di terminologia matematica).

Ci sembra che in questo equivoco ci sia una errata valutazione del significato della mossa che Frege compì al sorgere dell'indagine sui fondamenti. Sicuramente noi non conosciamo una filosofia

generale di Frege, ma sarebbe ingiusto dire che egli si è occupato di filosofia della matematica come se questa fosse una disciplina a sé, da indagare con metodi analoghi a quelli dello scienziato. Gli interrogativi di Frege avevano senso e determinatezza su uno sfondo filosofico preciso, quello per cui la distinzione fra analitico e sintetico è una distinzione profonda che riguarda fatti centrali della conoscenza umana: solo in questo contesto acquista la rilevanza che di fatto ha avuto il suo porsi il problema sulla natura della matematica. Tutto sommato nessuno ci obbliga a considerare come centrale la questione se l'aritmetica sia o meno analitica e anzi, per molti, questa non è una domanda particolarmente interessante ma per tutti è indubbiamente interessante la risposta di Frege, la «tecnica» che egli ha usato (questa sì valida in sé e per sé, elemento rilevante di una disciplina autonoma). In parole povere non esistono domande filosofiche «immediate», senza presupposti, sulla matematica, ma queste domande hanno senso su uno sfondo che è costituito da una ben articolata concezione filosofica, sia questa esplicita o meno. Possono essere gli strumenti invarianti rispetto agli sfondi, ma non le domande.

In questo senso a nostro parere non c'è un catalogo stabilito di problemi di fondo della filosofia della matematica e il crederlo può far correre il rischio o di prescindere dalla natura reale della matematica o di sussumere nell'ambito filosofico cattive formulazioni di veri problemi matematici. I recenti tentativi di coniugare indagini sugli assiomi forti dell'infinito e realismo matematico, sono un esempio istruttivo a questo riguardo.

Se quindi per il filosofo il problema è quello di adeguare sempre più le sue conoscenze alla matematica reale (al di là di tutte le ricostruzioni) su uno sfondo che deve rendere espliciti la matrice e il significato filosofico dei problemi che si pone, per il matematico la questione è diversa e in certo senso simmetrica: nel suo caso si tratta di trovare un linguaggio adeguato in cui esprimere quelle idee generali che alcuni, in verità non moltissimi, ritengono necessarie per una comprensione e un orientamento adeguati all'interno della loro disciplina. Abbiamo parlato di linguaggio perché ciò che rende spesso difficili da accettare le proposte filosofiche di molti matematici – pensiamo ad esempio agli stessi Hilbert o Brouwer ma venendo ai nostri giorni anche a MacLane o Lawvere o Thom – è che sono spesso espresse in un linguaggio che presenta

un singolare amalgama di richiamo all'efficienza, a criteri pratici di decisione (funziona/non funziona) e categorie filosofiche a volte collegate in modo arbitrario, senza tener conto che sullo sfondo esistono teorie filosofiche con una ben precisa struttura.

Qui, come nel caso del filosofo, il problema è la consapevolezza dello sfondo, è la possibilità di isolare problemi o categorie concettuali che cambiano significato a seconda del contesto. Il risultato è stato spesso una ambiguità in cui la convinzione secondo la quale le categorie filosofiche di fondo non si possono analizzare troppo in dettaglio, si combina con la fiducia veramente scientista che, una volta formulati come si deve, non esistano problemi insolubili, e che se questo non succede – allora – i problemi sono pseudoproblemi. Già Kreisel ricordava nel suo articolo *Hilbert's Programme* del 1958 come Hilbert esprimesse il problema della coerenza dell'aritmetica dicendo che bastava *solo* la convergenza di certi metodi di calcolo, quando in questo «solo» era racchiuso il raggiungimento di quello che era stato il sogno della sua vita scientifica. È questa ambiguità che fin dagli anni trenta portava persone come Gödel a non fidarsi del programma hilbertiano e a provare una forte antipatia per le posizioni di Brouwer, il cui richiamo filosofico era non poche volte un misto di propaganda e di filosofia di seconda mano.

Al fondo di questo atteggiamento c'è in molti casi la convinzione che in realtà *nessun* discorso sulla matematica ha un reale significato se non si traduce in indicazioni di carattere metodologico su *come fare* la matematica. Queste indicazioni possono prendere la forma di un programma fondazionale, come nel caso di Hilbert, che nella sua formulazione rinuncia non poche volte al rispetto per la complessità pur di raggiungere formulazioni il più possibile determinate sul piano operativo; oppure come nel caso di Brouwer possono divenire vere e proprie indicazioni per la costruzione di una matematica alternativa a quella data così che la metodologia si confonde con l'atto stesso di fare matematica. Questo tipo di esito lo ritroviamo anche in molti altri indirizzi di ricerca a noi più vicini e il fatto a nostro avviso ha un profondo significato.

Abbiamo già detto più sopra che all'interno della matematica è spesso difficile separare il dato dallo strumento, stabilire se una nozione descrive qualche cosa o semplicemente se è un modo per affrontare lo studio di un problema. Per prendere un esempio su

cui i dibattiti sono stati e rimangono tuttora molto accesi: l'idea del continuo come insieme di punti particolarmente strutturato è un dato intuitivo o comunque imposto dai fatti o è una ricostruzione strumentale che ci permette di affrontare problemi in cui compare il continuo? Gli esempi si potrebbero moltiplicare e con una certa dose di semplificazione ma tutto sommato realisticamente si potrebbe dire che mentre per il filosofo l'interesse per la matematica è spesso motivato in funzione descrittiva, viceversa per il matematico è in generale prioritario il carattere metodologico.

Si badi, questo anche se di fatto i discorsi filosofici sulla matematica riguardano nella maggior parte dei casi i suoi metodi dimostrativi e di concettualizzazione, più che i contenuti. Ma è appunto il carattere descrittivo di questi contenuti che pur non essendo quasi mai tematizzato dalla ricerca filosofica, viene quasi sempre *presupposto*. È a questo livello a nostro parere che può intervenire la logica, offrendo da una parte un linguaggio al matematico per tradurre le sue idee generali ed articolarle, dall'altra, per fornire uno strumento che permetta di misurare di volta in volta lo scarto fra descrizione e strumentazione, consentendo di indagare il mutuo intrecciarsi di questi due momenti. Non c'è nulla di particolarmente nuovo in questa tesi, perché da sempre la logica si è posta il problema di indagare la portata dei nostri strumenti conoscitivi. L'elemento di novità semmai è che nella sua veste di logica *matematica* essa può farlo in modo omogeneo alla pratica matematica stessa, potendo quindi trasformare le sue stesse analisi in strumenti e metodi matematici.

Sull'altro versante, la potenzialità della logica per l'indagine filosofica sta nella possibilità che offre di verificare le diverse tesi filosofiche su ciò che è strumentale e descrittivo all'interno della matematica. Tradizionalmente i filosofi hanno sempre cercato di separare ciò che è essenziale nella matematica da ciò che è tecnico come pure, nei casi in cui la matematica veniva applicata ad altre scienze, ad esempio la fisica, di distinguere tra ciò che è dovuto al linguaggio matematico e ciò che è sostanziale. Ora distinzioni di questo tipo, come abbiamo detto, non sono facili e richiedono analisi specifiche, valutazioni di analisi alternative, e in queste indagini la logica può avere un ruolo essenziale, mostrando come stessi temi possono richiedere strumentazioni matematiche differenti, come invece certe strumentazioni comportino ipotesi fattuali spe-

ifiche o viceversa ne siano determinate e così via. Quello che con-
ı è che in questo modo il filosofo non si trova più di fronte a una
nità monolitica da prendere o scartare in blocco, ma a qualche
osa su cui può fare delle sperimentazioni locali motivate comun-
ue, lo ripetiamo, da un essenziale sfondo filosofico che non è im-
licito nella matematica stessa.

Per tornare ad un problema che abbiamo ricordato sopra, quel-
ɔ del continuo, tutti noi sappiamo come all'interno di diverse
oncezioni filosofiche diverse idee sul continuo siano state presen-
ate, anche differenti da quelle su cui oggi si basa l'Analisi e le sue
pplicazioni alla fisica, ad esempio concezioni che usano gli infini-
esimi o che negano i punti a favore delle zone. Chi ha ragione?
Molto spesso i filosofi (tranne ovviamente quelli che non accetta-
ıo *tout court* il confronto con la matematica o con la scienza in ge-
ıerale) tendono a prendere per buono quello che la matematica
ıuale è praticata oggi fa o spesso addirittura, non potendo diretta-
nente attingere alla pratica matematica, quello che i matematici
licono di fare, e ritengono che in quanto non matematici non sia
ɔossibile opporsi alla scienza «del proprio tempo». Ci si riduce
:osì spesso a chiosare malamente quello che si crede sia la realtà
ɔggettiva dimenticando che su molti problemi di fondo, ancora ad
:sempio come quello del continuo, la matematica non ha una ri-
posta unica, si tratta di provare le alternative e saggiarne le po-
enzialità.

Ancora, è in gioco la consapevolezza che non è fissata una volta
ɔer tutte la separazione fra fatto e strumento e per giungere a di-
stinzioni di questo tipo, spesso essenziali per il filosofo, non c'è al-
:ra via che la sperimentazione alla luce di ben precise ipotesi. È in
ɿuesto senso che per noi probabilmente non esiste una filosofia
ɖella matematica come disciplina con precisi metodi e problemi:
ıon si tratta di rinnegare il contatto con la filosofia, come molti lo-
ɡici oggi tenderebbero a fare o riciclandosi come matematici puri
ɔ come informatici, col solo risultato in molti casi di perdere quel-
a visione d'insieme che sicuramente è stata una delle più profon-
ɖe e concrete motivazioni per il sorgere della logica matematica.
Se si perde questa prospettiva ci si chiuderà in un lavoro estrema-
mente specialistico che *somiglierà* a quello del matematico tradizio-
nale, ma che (ci si perdoni la franchezza) spesso non ne avrà il re-
spiro né la rilevanza.

Ci sono problemi filosofici nella matematica, come abbiamo cercato di dire, e questi sono o problemi che vedono nella matematica un test e una esemplificazione ma che devono avere una strumentazione e sfondo concettuale che trascendono la matematica stessa, oppure idee generali di carattere metodologico che cercano di portare alla luce tendenze del procedere matematico unificando, semplificando e saggiandone la validità. Tutto questo non costituisce una disciplina ma può essere sicuramente il terreno d'incontro fra discipline diverse quali la matematica, la linguistica, l'informatica ecc., un terreno in cui la logica ha uno suo preciso e fecondo ruolo autonomo di catalizzatore. È questa una conclusione che ci sembra emergere con naturalezza dalla nostra storia.

Dall'Ottocento a oggi, la logica formale è stata protagonista di un'evoluzione di straordinaria ricchezza e complessità. Da sempre alla confluenza tra filosofia e matematica, e nel nostro secolo anche tra linguistica e informatica, ne ha indubbiamente influenzato lo sviluppo. D'altro canto queste diverse discipline hanno spesso interagito con la logica, arricchendola con le loro problematiche.
L'ampia panoramica di Corrado Mangione e Silvio Bozzi offre una puntuale introduzione storica alla logica formale: dall'elaborazione di una matematica del pensiero (Boole) alla riflessione di Cantor sull'infinito, dagli ambiziosi tentativi di presentare la logica come sistemazione unitaria del pensiero astratto (con Frege) alla crisi dei fondamenti con l'insorgere dei paradossi (Russell), dalla revisione assiomatica di Hilbert ai teoremi di Gödel e al riaprirsi della contraddizione tra finito e infinito, tra astratto e meccanico; per concludere con l'emergere di logiche alternative (intuizionista, modale, polivalente, quantistica...) e di nuove forme di matematizzazione del mondo fisico (Analisi non standard), fino a presentare i nuovi problemi della complessità, precisati dall'avvento dei calcolatori elettronici.
Questo percorso viene qui esposto per la prima volta nella sua completezza e organicità, cercando di chiarire anche al lettore non specialista i nessi fondamentali dell'evoluzione della logica matematica e i suoi rapporti con il resto del sapere.

Corrado Mangione insegna Logica presso la Facoltà di Lettere e Filosofia dell'Università di Milano. Ha collaborato alla *Storia del pensiero scientifico e filosofico* di Ludovico Geymonat ed è autore di numerosi saggi e volumi di carattere logico-filosofico.

Silvio Bozzi, dopo aver insegnato per quindici anni presso il Dipartimento di Matematica dell'Università della Calabria, insegna attualmente Logica presso la Facoltà di Lettere e Filosofia dell'Università di Urbino. È autore di diversi saggi riguardanti le logiche non-classiche e la teoria dei modelli.

Grafica di Marco Volpati

ISBN 88-11-59966-0
9 788811 599661

L. 90.000
(prezzo di vendita al pubblico)